NONVERBAL COMMUNICATION IN INFANTS
AT-RISK FOR AN EVENTUAL DIAGNOSIS OF AUTISM

Shaye Benton Reavis

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Psychology (Clinical Psychology).

Chapel Hill
2007

Approved by:
Professor Gary B. Mesibov
Professor J. Steven Reznick
Professor Grace T. Baranek
Professor Martha J. Cox
Professor Scott L. Schwartz
ABSTRACT

SHAYE BENTON REAVIS: Nonverbal Communication in Infants At-Risk for an Eventual Diagnosis of Autism
(Under the direction of Gary B. Mesibov, Ph.D. and J. Steven Reznick, Ph.D.)

Little is known about the expression of Autism Spectrum Disorder (ASD) during infancy. Identification of autism in infancy is difficult because of a lack of diagnostic clarity for this age-group. Yet, research points to a variety of social and communication impairments in infants later diagnosed with autism. In this project, nonverbal communication delay, a potential early symptom of autism, was empirically investigated in a cohort of infants with behaviors that may place them at-risk for an eventual diagnosis of autism, and a comparison sample of typically-developing infants. Results indicated that there were statistically significant differences between groups in nonverbal communication, namely, for joint attention skills, for a subset of infants tested by one of the examiners. Differences between groups were present, though non-significant, when analyzing the entire sample. Both the entire sample and the subset showed some differences in their pattern of nonverbal communication, with the target group showing weaker joint attention than behavior requesting skills. Issues that may have influenced results, such as heterogeneity, timing and emergence of nonverbal communication development, and an examiner effect, are outlined. The relevance of the findings of group differences in nonverbal communication, particularly joint attention, is discussed.
ACKNOWLEDGEMENTS

There are many people who deserve special acknowledgement for their contribution towards the completion of this project. I would especially like to acknowledge my advisors at the University of North Carolina at Chapel Hill for their outstanding professional guidance, instruction, and training. Dr. Gary Mesibov has been my graduate advisor throughout my training in clinical psychology. Dr. Mesibov’s professional acumen and insights, extensive knowledge and experience, and longstanding mentorship and support have been primary in shaping my development throughout this project as well as my professional development in clinical psychology. Dr. Steven Reznick has also been my research mentor with whom I have worked for several years at UNC. Dr. Reznick’s proven excellence in research, his comprehensive and experienced knowledge-base, and his dedicated and supportive mentorship also have been essential in my development of research and academic skills and in my professional development as a psychologist.

Thanks is also extended to several individuals who participated in this project. Don Trull at the Frank Porter Graham Child Development Center provided hours of technological assistance, thoughtful advice, and the use of equipment. Chaka Coleman, Signe Boucher, and Nikki Henderson offered hours of their time in comprehensive assessments of children. Ndidi Okeke participated in several hours of reliability coding of the data. Dr. Cathy Zimmer of the Odum Institute at UNC Chapel Hill provided excellent statistical consulting support.

I would like to extend a thank-you especially to my parents for their thoughtful and
loving guidance, advice, and support, all of which have made possible my ability to succeed in graduate school and to complete this project. In addition, there are several other family members who have offered logistical and other essential support towards the completion of this project.

I want to thank most of all, my wonderful husband, Stephen, for his unending encouragement, support on all levels, and patience throughout all aspects of my graduate training and particularly during the completion of my dissertation. Finally, I especially want to acknowledge my sweet and wonderful daughter, Susannah, for her constant inspiration, and my darling infant daughter, Rebecca, born shortly after this project.
# TABLE OF CONTENTS

LIST OF TABLES ........................................................................................................................................ vii

LIST OF FIGURES ....................................................................................................................................... viii

Chapter

1. INTRODUCTION ......................................................................................................................................... 1

   Review of the Research on Early Characteristics of Autism ................................................................. 2

   Identification of Sample ......................................................................................................................... 15

   Examination of Nonverbal Communication Skills ............................................................................... 19

   Hypotheses .................................................................................................................................................. 26

2. METHOD ................................................................................................................................................ 27

   Design ....................................................................................................................................................... 27

   Participants ................................................................................................................................................. 27

   Procedure ..................................................................................................................................................... 31

3. RESULTS .................................................................................................................................................. 40

   Preliminary Analyses .............................................................................................................................. 40

   Core Analyses ............................................................................................................................................ 48
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Additional Analyses</td>
<td>53</td>
</tr>
<tr>
<td>4. DISCUSSION</td>
<td>60</td>
</tr>
<tr>
<td>Summary of Findings</td>
<td>60</td>
</tr>
<tr>
<td>Nonverbal Communication</td>
<td>65</td>
</tr>
<tr>
<td>Extreme Risk-Group Criteria</td>
<td>69</td>
</tr>
<tr>
<td>Examiner Effect</td>
<td>70</td>
</tr>
<tr>
<td>Autism and the Broader Profile</td>
<td>74</td>
</tr>
<tr>
<td>Limitations</td>
<td>77</td>
</tr>
<tr>
<td>Future Directions</td>
<td>78</td>
</tr>
<tr>
<td>Conclusion</td>
<td>80</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>101</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table

1. List of joint attention and behavior requesting nonverbal communication gestures coded during the Early Social Communication Scales (ESCS) assessment.......................................................... 82

2. Demographics of sample for Child Race, Maternal Education, and Birth Order... 83

3. List of independent variables and descriptions.................................................. 84

4. List of dependent variables and description ..................................................... 85

5. Linear regression models of explained variance in individual dependent variables by risk group; Original risk-group criteria.................................................. 86

6. Chi-square for surprise score differences between groups; Original risk-group criteria.......................................................... 87

7. Linear regression models of explained variance in individual dependent variables by risk group; Extreme risk-group criteria.................................................. 88

8. Chi-square for surprise score differences between groups; Extreme risk-group criteria.......................................................... 89

9. Paired sample t-tests comparing pairs of initiating and responding variables; Original risk-group criteria.......................................................... 90

10. Paired sample t-tests comparing pairs of initiating and responding variables; Extreme risk-group criteria.......................................................... 91

11. Examiner 1 cases only: Linear regression models of explained variance in individual dependent variables by risk group; Original risk-group criteria ....... 92

12. Examiner 1 cases only: Linear regression models of explained variance in individual dependent variables by risk group; Extreme risk-group criteria ....... 93

13. Paired sample t-tests comparing pairs of initiating and responding variables for Examiner 1 only; Original risk-group criteria.......................................................... 94

14. Paired sample t-tests comparing pairs of initiating and responding variables for Examiner 1 only; Extreme risk-group criteria .......................................................... 95
LIST OF FIGURES

Figure

1. Room Set-Up for Early Social Communication Scales (ESCS) .......................... 96
2. Scatterplot of Risk Score and IJA Composite Score delineating the 50th and 90th percentiles for children’s risk scores .......................................................... 97
3. Scatterplot of risk score and rate of total initiating behaviors .............................. 98
4. Bar graph for interaction effect of Examiner × Risk Group for the variable Rate of Total Frequency of Initiating Behaviors ..................................................... 99
CHAPTER 1
INTRODUCTION

Professionals have agreed for several decades that some symptoms of autism are apparent before 12 months of age (Rutter, 1978; Short & Schopler, 1988; Volkmar, Cohen, & Paul, 1986; Wing, 1980; Young & Brewer, 2002). Researchers who have analyzed home videotapes of children with autism find non-normative behaviors as early as 9 months of age (Baranek, 1999; Osterling, Dawson, & Munson, 2002). Parents of children with autism retrospectively report that they noticed “something was wrong” with their child during their child’s first year (Dahlgren & Gillberg, 1989; Short & Schopler, 1988). Despite these converging reports, autism is not yet regularly diagnosed in infancy. This is in part because DSM-IV-TR diagnostic criteria primarily apply to older children. In fact, several DSM-IV-TR items do not differentiate two year olds with autism from two year olds with other disabilities (Stone et al., 1999). These are:

1(b) failure to develop peer relationships appropriate to developmental level.

2(b) in individuals with adequate speech, marked impairment in the ability to initiate or sustain a conversation with others.

2(c) stereotyped and repetitive use of language or idiosyncratic language.

3(b) apparently inflexible adherence to specific, nonfunctional routines or rituals (DSM-IV-TR; American Psychiatric Association, 2000, pp. 70-71).

Another reason that autism is not diagnosed in infancy is that research has not yet identified
the early developmental trajectories of autism. The primary constraint in studying infants who show signs of autism is the difficulty in identifying these infants. Indeed, much of the research in this area has relied on retrospective reporting and coding methods. As a result, there is a lack of longitudinal research that begins in infancy (Young & Brewer, 2002).

This project provides a unique opportunity to study early behaviors that appear in infancy and may precede an autism diagnosis. By using a recently developed infant screening measure, a sample of infants who present with behaviors that may place them at-risk for an eventual diagnosis of autism will be identified. Secondly, the nonverbal communication skills of these infants will be assessed and compared to a control group of typically developing infants. As a backdrop, a review of what is currently known about characteristics of autism that may appear during infancy is first presented.

**Review of the Research on Early Characteristics of Autism**

Research that has examined early signs of autism has identified several salient characteristics and behaviors during the first two years of life. Current information is based on retrospective parent report (Dahlgren & Gillberg, 1989; Stone & Lamanek, 1990; Lord, 1995), retroactive analysis of home videotapes of first birthday parties (Baranek, 1999; Osterling et al., 2002), and prospective studies following children currently diagnosed with autism or children who have an older sibling with autism (Baron-Cohen et al., 1996; Gillberg et al., 1990; Suqiama & Abe, 1989). Information also has been obtained from the use of exploratory screening measures for children 18 – 36 months (Baron-Cohen, Allen, & Gillberg, 1992; Stone, Coonrod, & Ousley, 2000; Robins, Fein, Barton, & Green, 2001) and documentation of individual cases (Dawson, Osterling, Meltzoff, & Kuhl, 2000; Eriksson & de Chateau, 1992). The following summarizes the findings obtained from each approach.
**Parent Retrospective-Report Studies**

The earliest studies that examined early indicators of autism during infancy relied on parent retrospective-report. Researchers often conducted open-ended interviews with parents or asked them to select behaviors from a checklist that they remembered their child having. Parents who were included often had children who ranged widely in age at the time of the study, such as 4-27 years old (Dahlgren & Gillberg, 1989; Hoshino et al., 1982; Short & Schopler, 1988), indicating wide variability in how many years back parents were asked to recall. Despite the different reporting methods and variability in years of recall, parents consistently reported behaviors across these studies that fell into categories of social relationships, communication, play, imitation, and perception and sensory concerns (Dahlgren & Gillberg, 1989; Hoshino et al., 1982; Lord, 1995; Lord, Storoschuk, Rutter, & Pickles, 1993; Ornitz, Guthrie, & Farley, 1977; Short & Schopler, 1988; Stone, Lemanek, Fishel, Fernandez, & Altemeier, 1990). Furthermore, these reports differed significantly from retrospective reports by parents of children with mental retardation (Dahlgren & Gillberg, 1989; Short & Schopler, 1988; Hoshino et al., 1982), children with language disorders (Short & Schopler, 1988), and children with typical development (Dahlgren & Gillberg, 1989; Hoshino et al., 1982). Importantly, most parents reported that they noticed unusual behaviors by their child’s first birthday (Rogers, 2001).

Specific behaviors reported by parents across studies included “appeared isolated from his or her surroundings”, “had an empty gaze”, “had poor social interaction”, and “had abnormalities of gaze and response to sound” (Dahlgren & Gillberg, 1989). Other characteristics were “did not try to attract adult’s attention to own activity”, “did not follow mother with eyes”, “was overly responsive or nonresponsive to sensations” and “did not
orient to name” (Hoshino et al., 1982), which was also phrased as “did not respond to voice” (Lord, 1995). Short & Schopler (1988) found significant differences in items on the CARS as reported by parents of children later diagnosed with autism versus parents of children with mental retardation and those of children with typical development. Across these studies, the most frequently reported items included aberrant social behaviors (nonresponsiveness to voice or hearing one’s name, and preferring social isolation or having a tendency to be alone), and lack of communicative behaviors (failure to direct an adult’s attention to objects and failure to follow the gaze of another person). Sensory concerns were also sometimes reported by parents in these studies.

There are problems when using retrospective parent report. Parents’ memories of their children’s unusual behaviors as much as 15-20 years earlier could have been shaped over time by selective recall and by the way they have come to understand autism and their child’s development since that time. If so, retrospective parent report might not accurately reflect their child’s behaviors at an earlier age (Gillberg, 1989). Researchers can partially control for memory factors by including age-matched control groups of children with other disabilities for which symptoms were documented, or children with typical development (Dahlgren & Gillberg, 1989). Even with the inclusion of these control groups, researchers continue to find that children who later receive an autism diagnosis had early behaviors that set them apart from other children in their cohort and which may have been early indicators of autism. Furthermore, research that compares retrospective parent report of autistic symptoms during the preschool years and current assessment of symptoms in preschoolers who are already diagnosed with autism indicates that these two types of reports are largely congruent even when parents are asked to remember many years back (e.g., Lord, 1995;
Gillberg et al., 1990; Rogers, 2001).

Another influence on recall may be children’s current diagnostic status at the time of the study, whether that diagnosis is autism, mental retardation, or no diagnosis (Dahlgren & Gillberg, 1989). The behaviors that mothers report may relate more to what one would expect for the child’s current diagnostic status based on education and current information, rather than what the parent actually noticed at an early age. Fortunately, a large variability in behaviors is often reported by parents in these studies, which would not be expected if parents were reporting merely on the most typical autistic characteristics because of their child’s diagnostic status (Dahlgren & Gillberg, 1989).

In addition to problems with recall, external variables may influence parent awareness or understanding of problematic behaviors. Parent awareness of infant attributes may be influenced by education level, intelligence, birth order of their child, or mental health (Feldman & Reznick, 1996; Gillberg, 1989). It is difficult for researchers to interpret parent retrospective reports if individual differences based on these external influences are not taken into account. In addition, it is impossible to know how a parent’s baseline level of awareness or information influenced their report. Research has yet to determine how external variables may influence parent awareness and interpretation of their children’s behaviors. In summary, the types of behaviors remembered, the influence of diagnostic status, the present understanding of the child’s development, and the parents’ level of awareness or understanding of autistic behaviors may influence parents’ memories and subsequent report. Research that uses retrospective report often includes comparison groups to control for some of these factors, and although many behaviors across studies do converge, the difficulties in verifying parent report make these issues a concern nevertheless.
**Home Videotape Analysis**

More recently, researchers have observed and coded behaviors that may precede an autism diagnosis by viewing early home birthday video recordings of children later diagnosed with autism (e.g., Baranek, 1999; Osterling et al., 2002). Researchers often request videotapes of birthday parties because this better ensures that children are the same age when behaviors are coded, and it also provides a similar context across subjects. Typically, the mood is light-hearted, multiple adults are present, the child is the focus of attention, and ongoing conversation includes the birthday party per se.

The results from videotape analysis have consistently revealed that infants who will eventually have a diagnosis of autism tend to show unusual social behaviors, including failure to respond when their name is called, fewer looks at others or at visual stimuli, fewer social smiles, and fewer appropriate facial expressions; and delayed or absent communication, including fewer joint attention behaviors (e.g., pointing, showing, or alternating gaze between person and object), lack of affective expression, and lack of imitation when compared with typically-developing children (Adrien, Faure et al., 1991; Adrien, Lenoir et al., 1992; Adrien, Perrot, et al., 1992; Baranek, 1999; Losche, 1990; Mars, Mauk, & Dowrick, 1998; Osterling & Dawson, 1994; Osterling et al., 2002; Werner, Dawson, Osterling, & Dinno, 2000). Losche (1990) also found that children later diagnosed with autism had early difficulties with means-end behaviors, in that they became distracted by the process to the neglect of the goal, such as attending more to an object used to reach a goal than the goal itself (e.g., becoming distracted by a cup that needs to be moved out of the way to obtain a desirable present). Baranak (1999) also reported mouthing of objects and deficits in the areas of sensory-motor behaviors and aversion to touch, in addition to social
and communicative deficits. Across studies, researchers agree that the strongest differentiating factor when viewing and coding home videotapes of children later diagnosed with autism versus those of children with other disabilities or typical development is failure to respond to their name when called. This lack of response has been observed consistently, creating statistically significant differences between groups. Other behaviors that are repeatedly observed usually fall into social and communication areas of development.

Some retrospective video analysis studies have specifically examined behaviors that distinguished young children with autism from those with other developmental delays and mental retardation (Baranek, 1999; Osterling & Dawson, 1999; Osterling et al., 2002). Responding to spoken name, pointing, showing objects, and failure to look at people again were the strongest indicators of autism in retrospective home video studies by Osterling & Dawson (1999) and Osterling et al., (2002) in distinguishing children with autism from those with other types of disabilities. Baranek (1999) also compared children with autism to children with developmental delay and those with typical development. Again, responding to name continued to be the strongest discriminator between children with autism and the other two groups, in addition to orientating to visual stimuli, mouthing of objects, and displaying an aversion to social touch (Baranek, 1999).

In these studies, impaired language development, the presence of motor stereotypies, unusual motor posturing and slow development of gestures, all of which have been found to discriminate autism in older children, did not discriminate children later diagnosed with autism from those later diagnosed with other disabilities. Clinically, these behaviors were observed in the autism groups, but they did not reach the level of statistical significance when compared to children with other disabilities. Rather, it was the absence of negative
symptoms, particularly failure to respond to name or orient to speech and visual stimuli, including faces, that seemed to best discriminate the very young children (12 months or younger) from children with developmental disabilities and typically developing children.

There are methodological difficulties with retrospective videotape analysis. Using family home movies makes it difficult to ensure comparable taped segments in quality of recording and time involved. It is also difficult to control extraneous variables that may affect the child’s behaviors and interactions (e.g., number of people present, familiar versus unfamiliar people, how often adults initiate social interaction with the child, etc.) Another weakness is that studies have not often utilized control groups of children with mental retardation or language disorders to control for developmental problems nonspecific to autism (Osterling & Dawson, 1994). Finally, numbers of subjects have generally been smaller than in retrospective parent report studies, primarily because of the difficulties in obtaining first birthday or other home videotapes.

Prospective Studies

Prospective studies have been used to address some of the methodological difficulties with retrospective research. Baron-Cohen, et. al., conducted the most seminal prospective work, in which they followed 18 month olds who had failed key items on the Checklist for Autism in Toddlers (CHAT) (Baron-Cohen, Allen et al., 1992; Baron-Cohen, Cox et al., 1996). Baron-Cohen et al. (1992) followed 18-month-old children who had failed three items on the measure, including pointing to direct another’s attention to an object or event of interest, following another’s gaze, and pretend play. At a one-year follow up, these items continued to discriminate children with autism from typically-developing children. However, only when all three items were endorsed together did they differentiate children with autism
from developmentally-delayed children (Baron-Cohen et al., 1996). Protodeclarative pointing and pretend play alone did not discriminate children who received a diagnosis of autism from those who received a diagnosis of language delay. Furthermore, in a 5-year follow up, several children from the original sample who had passed these items were eventually diagnosed with autism. It is likely that delays in pointing to direct one’s attention and following gaze (deficits of communication) and pretend play are part of an early presentation of autism, but they are not the only early behaviors.

Gillberg et al. conducted another prospective study of children with autism to better discern developmental pathways of autistic behaviors (1990). These authors followed 28 children who had been referred for a possible autism diagnosis before the age of three, including 13 children who were less than 2 years old. The authors followed these children until they were at least three years old, finding that abnormal social development and unusual perceptual responses continued over time to be the strongest indicators of autism. Suqiyama & Abe (1989) conducted longitudinal work on 18-month-old children suspected of having autism, following them beyond 3 years of age. The authors found communication disturbances that prevailed throughout this time period in addition to some stereotypic behaviors. Looking at 20 month olds, Charman et al. (1998) found that lack of joint attention, poor or absent empathic response, delays in imitation, and aberrant play behaviors also distinguished toddlers with autism from those without (Charman et al., 1998). Similar to some of the home videotape analyses, Charman et al. failed to find that these behaviors discriminated 20 month olds with autism from those with developmental delay but no autism.

Recently, some researchers have followed younger siblings of children with autism prospectively to identify salient behavioral features that might indicate an eventual autism
diagnosis or characteristics that might be related to an autism spectrum disorder. Landa and Garrett-Mayer (2006) followed 87 infants, some who had an older sibling with autism and some who had no family history of autism, from 6-24 months of age. Children were later assessed with the Autism Diagnostic Observational Scale (ADOS) and classified following the study as unaffected, autistic spectrum disorder, or language delayed. The authors found that the differences between all three groups on language, motor, and visual reception became greater with age, with the ASD group performing significantly worse than either the language-delayed or the unaffected groups. Among other developmental variables, the authors found that receptive and expressive communication became increasingly divergent among all three groups as children got older, with differences first seen at the 14-month assessment.

Together, these prospective studies endorse the same general symptom categories of communication disturbance and difficulties with social interaction as differentiating autistic from typically-developing children. Unusual perceptual responses, poor imitation skills, and unusual play behaviors were also problematic for children with autism. These studies strengthen the idea that autism strongly affects communication and social abilities, but that it also permeates other areas of development including symbolic skills, pretend play, and sensory responses.

*Exploratory Screening Measures*

Several screening measures for autism in young children have been developed. Data gathered from the use of these measures supports results from retrospective and prospective studies, suggesting that the most frequent and consistent deficits fall within areas of communication and social ability. Stone et al., (2000) report on the Screening Tool for
Autism in Two Year Olds (STAT), which is an interactive screening tool designed for 24-35 month old children. The STAT examines aspects of play, imitation, and communication, including joint attention items. Research indicates that the STAT discriminates children with autism from those with developmental delays without autism, and from typical children, on the basis of joint attention, imitation, play, and requesting items.

The Diagnostic Interview for Social and Communication Disorders (DISCO), developed by Wing, Leekam, Libby, Gould, & Larcombe (2002), is an interview-based measure that allows a professional to assess autistic symptoms in children. Only some items on the DISCO have been shown to discriminate autistic from non-autistic groups. The items that most strongly discriminate the groups (p<.001) include lack of joint referencing/pointing, no interest in other children, inappropriate interactions, lack of emotionally expressive gestures, and failure to seek comfort when in pain or distress (Leekam, Libby, Wing, Gould, & Taylor, 2002).

The Autism Diagnostic Interview-Revised (ADI-R) is a semi-structured interview conducted with the parent (Lord, Rutter, & LeCouteur, 1994). The ADI-R has been shown to differentiate individuals with autism from those with mental retardation across all areas it assesses, including problems in social reciprocity, communication deficits, and restricted, repetitive and stereotyped behaviors. The ADI-R converges with the Autism Diagnostic Observation Schedule-Generic (ADOS-G) (Lord, 1997). The ADOS-G involves interacting with the child using several structured and semi-structured scenarios and assesses social interaction, communication, and play behaviors. Although the ADOS-G has not been examined for sensitivity and specificity, its nonverbal communication items, borrowed from an earlier version, the PL-ADOS (Prelinguistic ADOS), are able to discriminate autistic from
typical samples (Lord, 1997). Finally, the Childhood Autism Rating Scale (CARS) (Schopler, Reichler, DeVellis, & Daly, 1980) is a behavioral rating scale using play interaction or observation by the professional. Many items on the CARS have been shown to discriminate autistic from typical samples, including items pertaining to communication, social interaction, play behaviors, and imitation.

In summary, items on screening measures that differentiate children with autism from typical samples are consistent across studies and fall into categories of nonverbal communication, social interaction, and sensory behaviors. In the 3 to 4-year-old age group, when imitation and play skills should be well-developed, items regarding imitation and play also identify children with autism on screening measures.

*Case Studies*

Case studies that describe young children with autism generally concur with the findings of group studies. Authors describing case studies have noted difficulties in social interaction and poor eye contact, poor motor and vocal imitation, sleep dysregulation, motor stereotypies, and hypersensitivity (Dawson et al., 2000). Eriksson & de Chateau (1992) describe a child who at 13 months became withdrawn from her surroundings, had poor babbling, often failed to respond to her name, and sat alone with little interest in playing with others, instead showing solitary, stereotyped and routine play. These notes were drawn from her medical records during pediatrician visits and parent report. Again, these specific behaviors can be classified within communication and social categories. Dawson et al., (2000) describe a young child who showed signs of poor regulatory behaviors, poor eye contact, lack of imitation abilities, and sensorimotor difficulties before the age of two. This child later received an autism diagnosis. The early chief concerns in this case involved
sensorimotor and regulation behaviors.

Summary of Research on Early Characteristics in Autism

An accumulation of data indicates that children who are diagnosed with autism show problematic behaviors as early as 12 months of life, and possibly sometime during the first year. Findings converge across every research approach that there are early disturbances in areas of communication and social behaviors for both younger and older children with an autism diagnosis. Studies that have targeted younger children (12-24 month old) also report sensory processing and self-regulation problems, whereas research that focuses on older children (3-4 years) often reports imitation and play deficits. In particular, specific behaviors that emerge across studies that apply to very young children include failure to orient to name, withdrawal from social interaction, and failure to develop nonverbal communication skills appropriate to age.

Nonverbal communication stands out as a primary deficit during infancy for children later diagnosed with autism. Across multiple studies, parents and researchers report delays across many types of nonverbal communication behaviors, including gesture use, gaze following, eye contact, responding to others’ communicative bids, and initiating communication with others. However, there are several aspects of these deficits that are not understood. Little is known about the extent of delay for nonverbal communication in infants who may later receive an autism diagnosis. The interrelationships among types of nonverbal communication also have not been studied for infants who may have autism. Therefore, an important step in the study of early indicators of autism is to examine the pattern of nonverbal communication deficits as they appear during infancy.

This research review highlights the importance of nonverbal communication as an
area for additional study, but it also identifies difficulties in carrying out such research with infants. First, much of the research conducted to date has been retrospective. There is a need for identifying infants who currently present with delayed behaviors that may suggest the early manifestation of autism. Although several screening measures for autism exist, most of them are designed for toddlers and older children. In order to study deficits of nonverbal communication in infants, screening measures need to identify infants with problematic developmental patterns consistent with autism.

This project presents a unique opportunity to examine nonverbal communication deficits in infants. In order to study nonverbal communication, a sample of 12-month-old infants who currently present with a pattern of behaviors that are believed to precede an eventual diagnosis of autism will first be identified using a recently developed and empirically-based early screening measure. Although the early behavior patterns and developmental profile for infants who may later receive an autism diagnosis is not completely known, the group of infants identified for this study comprise a sample of infants who present with a behavior pattern that is the “best-guess” to date, based on the research literature, as to what behavioral constellation and early characteristics infants with autism might have. These infants are not diagnosed with autism based on this measure, given that a diagnosis cannot be made at this age, and the children will not be receiving a formal clinical assessment, given that there is not enough knowledge at the present time to conduct such an assessment. However, based on the assumption that these infants have a pattern of behaviors that suggest eventual autism, their nonverbal communication skills will be examined. First, the identification of the sample to be used in this study will be outlined because it is an important and unique aspect of this research.
Identification of Sample

In order to identify a sample of infants who have a pattern of behaviors consistent with the early manifestation of autism, a recently developed infant screening measure, the First-Year Inventory (FYI) (Baranek, Watson, Crais, & Reznick, 2004; Reznick, Baranek, Reavis, Watson, & Crais, in press) was mailed to parents of 12 month olds in the Raleigh/Durham/Chapel Hill area. The FYI will be described in detail because the identification of the sample is a critical component of this research.

The First-Year Inventory (FYI)

The FYI is a 63-item checklist for parents to complete when their child is 12 months old. The measure asks about a range of behaviors across social, communicative, sensory, and motor developmental areas. Items include behaviors typically seen in 12-month-old infants that, when present to varying degrees or in certain combinations, may presage the emergence of an autism spectrum disorder. For the first 46 items, parents select the frequency that they believe their child displays various behaviors on a 4-point scale from “Never” to “Often”. Fourteen additional items ask parents to complete a sentence with one of four options that most closely describes their child. One item asks parents to circle letters their baby can say, and two items ask about developmental, physical, or medical concerns. The FYI takes approximately 10-15 minutes for the parent to complete.

The FYI was developed by a team of researchers specializing in early development and autism. The items are based on behaviors derived from a comprehensive literature review, which yielded 26 original behavioral and developmental categories. This review covered empirically-based retrospective, prospective, and longitudinal studies as well as case studies that reported characteristics that were noticed and recorded by parents and
professionals for young children who eventually received a diagnosis of ASD. The 26 categories covered communicative, social, affective, and sensory concerns and included protodeclarative pointing, gaze monitoring, hearing, lack of stranger anxiety, sleep, feeding, and regulation, autistic aloneness, and poor social interaction. Similar items were collapsed into single categories. Items were written using the categories as a guide. The wording of items was determined by group discussion and expert consensus.

Several versions of the FYI were produced. Each version was submitted for feedback from experts in the field of autism and parents. The authors used this feedback to modify the FYI for clarity, to ensure that all desired behavioral areas were being targeted and that each item had at least one response alternative that was notably atypical. Three pilot mailings were conducted, from which parent responses were used to revise and reword items that had no discriminatory power (i.e., had no response alternative that was unusual), or that were misunderstood (i.e., most parents left blank or failed to endorse the intended most typical response).

A validity study was conducted using the FYI as a retrospective measure (FYI-Retrospective Version, or FYI-R) (Watson et al., 2007). In this study, the construct validity of the FYI-R was examined by asking parents of preschoolers to answer the FYI-R items based on what they remembered about their child’s behavior at 1 year of age. Responses were then compared for parents of preschoolers diagnosed with autism (n=38), parents of preschoolers diagnosed with a developmental disability but no autism (n=15), and parents of typically-developing children (n=40). FYI-R risk scores of children with autism were significantly higher than either children with developmental delays or typically-developing children. The developmentally-delayed children also received significantly higher risk scores.
than the typically-developing children. Results also indicated that the FYI-R had high
sensitivity in detecting children with autism when the criterion for risk was set at the 90th
percentile, but at this level the FYI-R also detected many children with developmental
disabilities, and therefore had lower specificity. When the criterion for risk was set at the 98th
percentile, the specificity of the FYI improved, with zero children with developmental
disabilities receiving a risk score this high. However, the sensitivity of the measure
decreased, as several children with diagnoses of autism had also received risk scores that fell
below the 98th percentile.

FYI Scoring

The development of scoring of the FYI underwent several stages. In an initial
approach, infrequent response alternatives were assigned either 1 or 2 risk points depending
upon exactly how infrequently the response alternative was selected by parents in the pilot
mailings. Children who accrued a higher total number of risk points were considered to be at
higher risk. This approach to scoring was straightforward but was dramatically affected by
the distribution of types of questions included in the FYI.

A more sophisticated scoring of the FYI takes underlying constructs into account. A
large sample of normative data was analyzed using a process of “construct-shaping”, in
which factor analysis and then item-total correlations were used to sort the FYI questions into
thematically-related separate constructs (see FYI manuscript, Reznick et al., 2006 for greater
detail about this approach). This shaping process yielded a total of 8 constructs, which
include: Social Orienting and Receptive Communication; Social-Affective Engagement;
Imitation; Expressive Communication; Sensory Processing; Regulatory Patterns; Reactivity;
and Repetitive Behavior. A factor analysis on all eight factors revealed two distinct domains.
The first domains represented the first four constructs listed above and was labeled “Social-Communication Behaviors”. The second domains included the second four constructs listed and was labeled “Sensory-Regulatory Functions”.

A total risk score for each child is calculated in the following way. Risk scores are first tallied across questions for each of the eight constructs. For each question, the least frequent response alternative is assigned one risk point and responses that are extremely unusual (selected by less than 5% of all parents) are assigned a second risk point. A final risk score is generated using an algorithm that translates the total number of risk points for each construct into a single risk score ranging from 0 to 50 (with the 25th percentile being equivalent to a score of 2.4, the 50th percentile being equivalent to a score of 5.5, the 75th percentile being equivalent to a score of 10.0, the 90th percentile being equivalent to a score of 14.2, and the 98th percentile and above being equivalent to a score of 20.7 and greater). This algorithm is based on a quasi-logarithmic scale that awards more points to children who have a greater number of unusual answers. The final risk score can then be calculated by averaging risk points across the eight constructs.

**Present Purpose of the FYI**

Children with high FYI scores do not “have autism” in any clinical sense. The primary deduction is that they have a large number of behaviors and characteristics that may predict an eventual diagnosis of autism. Projects that follow infants longitudinally will eventually determine whether a high FYI score is indeed a risk indicator of an eventual diagnosis of autism. As more is learned about early atypical behaviors, it may be possible someday to assign a diagnosis of autism spectrum disorder by the first birthday, if not earlier. In the present study, the FYI was used to identify a sample of infants who presented with a
pattern of behaviors that are believed, based on face validity and retrospective predictive
validity, to indicate atypical development that is consistent with clinical expectations about
the early stages of autism spectrum disorder and an eventual diagnosis.

The present study is based on a comparison of two cohorts of infants. One cohort, the
target group, was selected based on their high FYI scores (80th percentile and above,
discussed under Methods), indicating that their parents reported a large number of behaviors
that may place them at-risk for an eventual diagnosis of autism. This target group was
compared to a control group of infants who were matched to the target group and selected
based on receiving average FYI scores.

Examination of Nonverbal Communication Skills

The second stage of this project was to assess nonverbal communication skills, one of
the strongest deficit areas suggested by the literature for young children with autism, in the
sample of infants identified using the FYI. In order to better understand the extent and pattern
of nonverbal communication delay in 1-year-olds at-risk for autism, the pattern of nonverbal
communication skills was compared between the target and comparison samples.

Research has documented nonverbal communication delays for preschool and school-
aged children with autism who score lower on measures of nonverbal communication than
age-matched, typically developing peers and mentally-aged matched peers with
developmental disabilities without autism. As noted earlier, nonverbal communication delays
are also one of the most salient features in infants and young children ages 12-24 months
who later receive an autism diagnosis. (Baranek, 1999; Dahlgren & Gillberg, 1989; Osterling
et al., 2002; Stone et al., 2000). For example, at least one study using home videotape
analysis and at least one prospective study comparing children with autism, children with
general developmental delay, and children with typical development showed that a type of nonverbal communication known as joint attention discriminated children with autism and developmental delay from the typical group, but not children with autism from children with developmental delay (Baranek, 1999; Charman et al., 1998; Osterling & Dawson, 1999; Osterling et al., 2002).

Although studies such as these have documented nonverbal communication deficits in infants, it is not known exactly how delayed nonverbal communication skills are for these infants relative to infants who are developing along a typical trajectory. This question will be addressed by assessing nonverbal communication patterns in the proposed cohorts. First, because nonverbal communication is delayed in older children with autism, it is predicted also to be delayed at 12 months of age in infants who are at risk for an eventual diagnosis of autism, when many aspects of nonverbal communication are emerging in typical development. Therefore, the first hypothesis for this project is that target children will perform significantly lower on measures of nonverbal communication relative to the comparison sample.

In addition to showing global delays in nonverbal communication, children at least three years old and older have been shown to have an autistic-specific pattern of nonverbal communication delay, in that they develop nonverbal communication for nonsocial purposes first, and then later develop nonverbal communication for social purposes, with their social nonverbal communication skills lagging far behind on a permanent basis (Mundy, Sigman, & Kasari, 1994; Mundy, Sigman, Ungerer, and Sherman, 1986; Sigman, Mundy, Sherman, & Ungerer, 1986). Thus, children with autism tend to have stronger abilities in nonverbal communication used for nonsocial purposes. For example, they are able to extend their arm...
with an open hand to communicate that they want a colorful toy, give their cup to an adult to indicate that they are finished drinking, or make eye contact with an adult to communicate that they want a cookie. In contrast, children with autism tend to have weaker abilities in nonverbal communication used for social purposes. For instance, they are less likely to point to an interesting picture to express their delight to a caregiver, show their favorite toy to an adult, or look at a parent and laugh when they see a puppy. These nonverbal communicative gestures that have a social function are known as joint attention behaviors. Nonverbal communicative gestures that have a nonsocial function are known as requesting behaviors. These two types of nonverbal communication are described in more detail below, along with what is known about them in the autism population.

**Social Nonverbal Communication: Joint Attention Behaviors**

Joint attention involves the ability to coordinate attention with another person through the use of eye contact and gestures, in order to share the affective or social experience of an object or event (Bruner & Sherwood, 1983; Seibert, Hogan, & Mundy, 1982). Behaviors early in development that reflect this ability to coordinate attention include *referential looking*, in which a child looks back and forth between the caregiver and objects to share enjoyment of that object; *protodeclarative pointing*, in which a child points to show the caregiver an object, person, or event; and *showing*, in which a child hands an object to the caregiver to draw her attention to that object (Bakeman & Adamson, 1984; Butterworth, 1991; Corkum & Moore, 1995; Osterling et al., 2002; Scaife & Bruner, 1975). A distinction is also usually made between initiating and responding joint attention skills. Thus, if a child points to a toy to direct the parent’s attention to that toy, the child is engaged in initiating joint attention. If, however, the parent points to a toy to show it to the child and the child
responds by following the point and looking at the toy, the child is engaged in responding joint attention. Children can initiate or respond to joint attention eye contact, pointing, or showing behaviors. These early joint attention behaviors typically emerge from 9 to 18 months of age, as infants gradually learn to coordinate their attention and actions on objects with the attention and actions of other persons (Bakeman & Adamson, 1984; Scaife & Bruner, 1975).

Joint attention is distinguished from other infant looking-behaviors by several key aspects. First, joint attention is deliberate and communicative unlike coincidental eye contact that occurs between mother and infant from birth (Tomasello, 1995). Joint attention occurs when a child seeks out another person’s attention through visual and nonverbal gestures, in order to direct the person’s attention to a third object or event. Hence, the infant learns how to actively influence another person’s attentional focus. Second, joint attention involves the expression of affect, most commonly surprise, pleasure, or fear. This affect must be ‘joint’, or shared, such that infant and caregiver seem to realize that the other is experiencing the same emotion.

Third, a child learns how to include a third party in the communicative exchange. As joint attention ability emerges, early dyadic attention-sharing, which involved parent and infant only, becomes a triadic attention-system, which involves parent, infant, and the object or event upon which they attend (Dunham & Moore, 1995). These three aspects of joint attention (intentionality, affective exchange, and shared awareness for a third party) make joint attention a foundational skill by which the infant develops social and affective intelligence. With the development of joint attention, the infant opens a major channel for learning about the world, understanding how others connect affect with experience, and
developing an increasingly complex understanding of social, affective, and informational exchange. This foundational skill becomes increasingly complex as the child grows older. Children learn to share attention with others not only for objects, but also for situations, abstract ideas, complex emotions, and subtleties of communicative exchange and symbolic meaning.

Extensive research documents that children with autism are quite delayed in the development of joint attention in comparison to same-aged peers (Landry & Loveland, 1988; Loveland & Landry, 1986; Mundy et al., 1986; Mundy et al., 1994; Sigman et al., 1986; Wetherby, 1986). Results from various studies have found that measures of joint attention alone can identify 73-94% of preschool-aged children with autism from children with mental retardation matched on mental age, or from children with other developmental disabilities (Leyw & Dawson, 1992; Mundy et al., 1986; Mundy et al., 1994; Mundy, Sigman, & Kasari, 1990). A large body of research shows that children with autism are impaired in many aspects of joint attention, including lower levels of initiating and responding to referential looking, protodeclarative pointing, and showing (Mundy et al., 1986; Mundy et al., 1994; Mundy, 1995). In addition, research has shown that children with autism are usually more impaired in initiating joint attention versus responding joint attention.

Research shows that many higher-functioning children with autism eventually develop some joint attention skills (Dilavore & Lord, 1995; Mundy et al., 1994). Higher-functioning 5- and 6-year-old children with autism can learn to respond to joint attention bids from caregivers, although they often lack the initiative to enact these behaviors (Dilavore & Lord, 1995; Mundy et al., 1994). The majority of individuals with autism continue to show joint attention and social deficits throughout the lifespan, although studies estimate that 17-
47% of individuals with autism improve on some social skills in adolescence and adulthood. Individuals with higher IQs and more educational and intervention opportunities show the most improvement (Gillberg & Steffenburg, 1987; Kobayashi, Murata, & Yoshinaga, 1992).

**Nonsocial Nonverbal Communication: Requesting Behaviors**

Another set of nonverbal gestures are used for making requests and are known as behavior requesting skills. Communicative actions and gestures similar to those used in joint attention (e.g. eye contact, reaching) are observed in children with autism, although most researchers agree that these behaviors are more often used for instrumental purposes, such as requesting or refusing an object, rather than sharing affect about an object or experience (Baron-Cohen, 1989; Curcio, 1978; Mundy et al., 1986; Mundy et al., 1990). Children with autism show these actions that appear the same on a descriptive level (e.g., reaching out their hand, making eye contact). However, these children fail to show nonverbal behaviors when the function is to share affect about a third party, object, or event. Children with autism do not always demonstrate an understanding that communicative partners have the status of personhood. Instead, they often use visual and gestural communication to treat the communicative partner as an instrumental agent to obtain a goal, rather than a social, interactive partner with whom they can share experiences (Curcio, 1978). For instance, children with autism might request an object by reaching and opening their palm, looking at their parents, or holding out an object for another person to put away. Similar to joint attention, children can also either initiate behavior requests or respond to behavior requests. For instance, if a child points to obtain an object, the child is engaged in initiating behavior requesting. On the other hand, if a parent holds out her hand to request that a child gives her a toy and the child complies, the child is engaging in responding behavior requesting.
Most studies on joint attention and requesting behaviors have examined the relative frequencies of these skills at slightly older ages, when children are more eligible for a diagnosis (around 2-3 years of age). At these ages, children tend to show a greater frequency of behavior requests and very little joint attention. It can be assumed that because these two types of nonverbal communication are so diverging for young children with autism, they also must develop at different rates at 9-12 months of age as well. It is predicted that this splintered pattern of nonverbal communication abilities will be found in 12 month olds whose FYI profiles indicate characteristics believed to precede an autism diagnosis, and that the relationship between the two skills will differ markedly from 12 month olds with a typical behavioral profile. **The second hypothesis for this project is that the target sample of children will show a deviant pattern of nonverbal communication, with stronger abilities in requesting gestures and weaker joint attention skills, relative to the comparison sample, which will show similar frequencies of these two types of behaviors.**

**Measurement of Nonverbal Communication**

Two measures will be used to assess nonverbal communication. The first, the Early Social Communication Scales (ESCS), was designed by Siebert, Hogan, & Mundy (1982) to assess the nonverbal communication skills of joint attention and requesting behaviors. This measure is psychometrically sound and widely used to assess patterns of nonverbal communication (Seibert, Hogan, & Mundy, 1982; Delgado et al., 2002). The assessment follows a loosely structured set of interactive play activities that encourage the child to use nonverbal communication. Extensive coding guidelines define various nonverbal communicative behaviors across the various play tasks. Codes are based on the presumed function of the child’s nonverbal behaviors, which allow differentiation of the child’s social
(joint attention behaviors) versus nonsocial (requesting behaviors) intent. The second approach to measuring joint attention and requesting is a set of tasks adapted from several different authors and combined to target joint attention and requesting skills. These tasks take much less time to administer relative to the ESCS but provide a quick and easy opportunity to elicit initiating joint attention and initiating behavior requesting bids, with the final task providing the child with an opportunity to demonstrate a nonverbal response to three joint attention bids by the examiner. These tasks are described in more detail under Measures.

By using multiple assessment techniques, convergent measurement can be used to assess the constructs of nonverbal joint attention and requesting skills. This is preferable to a single assessment tool because using only one measure provides less confidence in construct validity and makes no allowance for measurement error. Any single measure yields error variance, but by averaging across multiple measures, random error variance can be reduced.

Hypotheses

In summary, the hypotheses for this project are as follows:

1. Children in the target sample will perform significantly lower on measures of nonverbal communication relative to the comparison sample. Specifically, children in the target group will show lower scores on both joint attention and requesting skills relative to comparison children.

2. Children in the target sample will show a deviant pattern of nonverbal communication, with stronger abilities in nonverbal requesting behaviors and weaker joint attention skills, relative to the comparison sample, which will show similar frequencies of the two types of behaviors.
CHAPTER 2

METHOD

Design

This study has been designed using a passive observational approach in its comparison of two groups of children who present with different FYI scores. The goal of the study is to make inferences about observed differences in developmental variables, particularly communication patterns, between the two groups of children. The purpose of making these comparisons is to attempt to identify communication patterns that may be associated with a broad presentation of characteristics in 1-year-olds that may indicate the early manifestation of autism.

Participants

FYI Mailings

For the first phase of this project, FYI screening measures were mailed to families during a 9-month period of time. Families’ names were obtained from birth certificates on file with the North Carolina Department of Vital Records, which comprise a database developed by Dr. Steven Reznick. Families were selected to receive a mailing if they lived in zip code zones that were generally within 30 miles of Chapel Hill, and if they had an infant who was 1 week less than 1 year at the time of the mailing. Families were excluded if they self-identified on the birth certificate as Spanish speaking, due to the fact that the FYI has not yet been translated into Spanish. Research does not indicate that infants in Spanish-speaking families should be any more or less at risk for autism than infants in English-speaking homes,
nor should their development be different on the relevant dimensions. In fact, cross-cultural studies report remarkable consistency of the prevalence, intellectual abilities, gender differences, and social class factors associated with autism (Klinger & Dawson, 1996). Although the exact number of families who did not participate due to language is not known, it is estimated that about one-third of the total surrounding population is Spanish-speaking.

FYI packets were mailed to a total of 6,304 families during 2004-2005. Envelopes were returned for 363 incorrect addresses. Out of the 5,941 FYI mailings that were successfully delivered, 1,496, or 25%, were returned, with higher return rates for Caucasian families and those with greater educational attainment. Forms were eliminated from 39 families due to parents filling them out past the 4-week requested time period of their child’s birth date, and from 166 families due to the child being born prematurely (prior to 37 weeks gestation). Nine forms were both late and also pertained to premature infants, resulting in a total of 196 forms being eliminated. The final sample of FYI forms totaled 1300. Out of these, approximately 95% of parents provided consent to be contacted for lab assessments.

Recruitment for this study

Participants for this study were drawn from the 1300 returned and completed forms. Children were selected for the risk group on the basis of their FYI scores, and children were chosen for the comparison group based on their ability to be matched with children in the risk group on gender, ethnicity, maternal education, and birth order. Children were excluded from the study if parents reported a serious medical condition or genetic diagnosis (e.g., prematurity, cerebral palsy, cystic fibrosis, Down syndrome, fragile X syndrome).

The 80th percentile was selected as the best division of the data into risk versus non-risk groups, based on the total distribution of FYI data accumulated at the time. An FYI score
at the 80th percentile indicated that the parent was reporting many behaviors that represented the least frequently reported and most unusual behaviors relative to all parents who had completed the form. Based on the scoring of the FYI and the distribution of data collected at the time, an FYI score of 9 was equivalent to the 80th percentile. Therefore, children were initially selected for the risk group if they had an FYI score of 9 or higher. After the recruitment and assessment of the sample for this study, the FYI scoring was modified based on additional FYI data. This change in scoring resulted in slightly modified risk scores for children in this study, with the 80th percentile now being equivalent to an FYI score of 11 instead of 9. The new scoring then resulted in a slight shift of the risk-groups for this sample, with six children in the target group now qualifying for the comparison group.

Recruitment of the final sample for this study was conducted as follows. Each month during the 9-month period of this study, a research assistant contacted all families whose child had received an FYI score at the 80th percentile or above and who had provided consent to be contacted again. When a family agreed to volunteer for the lab portion of the testing, the research assistant then contacted potential families whose child had received a risk score below the 80th percentile and who matched with the target child on the criteria listed above. This procedure in general resulted in approximately 6-8 families who volunteered each month for the lab-based assessments.

Some limitations made it difficult to recruit more families, including cancellations during the winter due to child illnesses and inclement weather; testing was not conducted around holidays and exam time due to the families’ commitments and the examiners’ school schedules; finally, each examiner was able to test a maximum number of children each week (3 children for one examiner and 1 child for the second primary examiner). The resulting
number of families who participated in this study totaled 56. Based on the old scoring, 31 of these children were recruited for the target group and 25 for the comparison group; with the new scoring, 25 children met criteria for the target group with 31 in the comparison group. Because the new scoring reflected additional knowledge and information about the FYI, groups were divided and analyses were run based on the new scoring.

Avoiding experimenter bias

A research assistant who did not conduct the testing was given the task of assigning families to examiners and scheduling the appointments. Several steps were taken to ensure that the research assistant did not introduce bias in the assignment of families to examiners, and that examiners remained blind to each child’s group status. First, FYI forms were identified only by the child ID number and not by parent or child name. These identification numbers linked each FYI to its matching Subsequent Participation Form, which included the family contact information. When the research assistant received completed FYI packets, she separated the FYIs and Subsequent Participation Forms into two piles. She then selected children for the target and comparison groups based on their total FYI scores, gender, ethnicity, maternal education, and birth order. Following this selection, the research assistant mixed the ID numbers alone and then matched the IDs to the parent names on the Subsequent Participation Forms. She then used the contact information from the Subsequent Participation Forms to schedule parents for the assessments. The research assistant assigned children to one of three examiners for the lab assessments. In this way, the research assistant did not know which children belonged to which sample group during assignment and scheduling.
Procedure

Measures

The Early Social Communication Scales (ESCS). (Seibert et al., 1982). The ESCS is designed to elicit and assess the frequency and function of two nonverbal communication skills – joint attention and requesting behaviors. This measure is typically used with children whose mental ages range from 8 to 30 months and takes 15 to 25 minutes to administer. The ESCS is administered using several toys that are placed next to the examiner, in view but out of reach of the child. The child sits in the caregiver’s lap, with the examiner directly in front on the other side of a table. Toys include 3 small wind-up gadgets, 2 hand-held mechanical toys, a hat, a comb, a book, a ball, a car, an uninflated balloon, and a closed jar filled with cheerios. Three colorful posters adorn the walls behind and to the side of the child. See Figure 1 for the positioning of the examiner, child, and toys.

The examiner presents and/or activates one toy at a time, keeping the toys beyond the reach of the child. Intermittently, the examiner points to the wall posters. The procedure is structured so that the same toys and a consistent number of trials are given per child. The ESCS has some standardized instructions and gestures for the examiner to follow per trial (e.g., to request a toy from the child, the examiner says “Give it to me” with open palm, spoken three times; to elicit a head-turn, the examiner points at a wall poster with her elbow on the table and eyes directed towards the poster, away from the child). This standardization ensures that each child has the same opportunities to initiate or respond with nonverbal communicative gestures. The ESCS manual also emphasizes the importance of maintaining ecological validity and encourages the examiner to modify aspects of the assessment, such as the order of toy presentation, to elicit the best interaction and behavioral performance of the
Because the ESCS assesses nonverbal communication, child-adult interactions throughout the ESCS are primarily nonverbal. The examiner speaks as little as possible, refraining from speech when toys are presented to avoid triggering a response from the child based on the examiner’s verbal initiation. The procedure is videotaped for later coding that follows a standardized coding scheme for frequency, type, and function of behavior. Nonverbal child gestures that are initiated by the child are coded according to type (eye contact, pointing, reaching, giving, etc.) and classified by their function as either joint attention or requesting behaviors. Certain tasks are also designed to assess responding joint attention and responding behavior requesting. Children receive frequency scores and, for certain tasks, total correct scores for: (a) initiating joint attention; (b) responding joint attention; (c) initiating behavior requesting; and (d) responding behavior requesting. The variables derived are continuous. The ESCS has solid psychometric properties (Seibert et al., 1982) and high levels of interrater reliability for functional category scores (McEvoy, Rogers, & Pennington, 1993). See Table 1 for a list of joint attention and behavior requesting gestures coded by the ESCS for this project.

Combined Set of Nonverbal Communication Tasks: The Surprise Task and Car Task (Butterworth & Adamson-Mecado, 1987; Butterworth & Jarrett, 1991; Lord, 1997; Scaife & Bruner, 1975). These tasks comprise a set of loosely structured interactive tasks between the child and examiner. Each task has been used repeatedly in prior research as a measure of nonverbal communication skills. These tasks have been combined into two main tasks and

---

1 The ESCS also includes tasks that are coded for social interaction behaviors, such as turn-taking, teasing, and social play, although these scores were not analyzed for this project.
have been named, respectively, the Surprise Task and the Car Task.

The Surprise Task consists of two simple tasks. The first task is drawn from Scaife and Bruner’s (1975) prototypical joint attention paradigm, in which the experimenter plays with the infant, suddenly stops, and turns her head in silence towards a prespecified direction. A child who follows the head turn with his or her own head turn is presumed to be manifesting joint attention. The second task, adapted from Butterworth and Jarrett (1991) adds other communicative cues to the gaze-monitoring task. When the experimenter turns her head, she looks at a predetermined, brightly colored shape on the wall and then gasps as though surprised by what she sees. The experimenter then enacts this behavior again with another shape on the wall, this time also pointing to the shape in addition to turning her head and gasping. The Surprise Task (the combination of these two tasks) assesses the child’s ability to respond to joint attention. For the Surprise Task, the child receives one categorical score indicating how many behavioral cues are required for the child to respond to the examiner’s joint attention bid (1=child needs only the examiner’s head turn in order to respond; 2=child needs the examiner’s head turn and surprised expression in order to respond; 3=child needs the examiner’s head turn, surprised expression, and point in order to respond; 0=child does not respond to any of the behavioral cues).

The Car Task consists of two simple tasks developed and used by previous investigators to assess aspects of nonverbal communication. The first task was developed by Butterworth and Adamson-Mecado (1987). Here, the experimenter enables a remote controlled toy (e.g., a car) that has been placed one to two yards from the child. The toy is designed to elicit joint attention and behavior requesting eye contact, reaching, and pointing. Second, in a paradigm adapted from the ADOS-G (Lord, 1997), the experimenter places the
remote controlled toy closer to the child but out of reach. Finally, the toy is placed in the child’s hands. Each child is allowed to view or hold the car during all three steps for the same amount of time. For each part of this task, initiating joint attention and initiating behavior requesting, including eye contact, reaching, pointing, showing, and giving, are coded using the same coding scheme developed for the ESCS. For the Car Task, the child receives one continuous variable score for total initiating joint attention bids and one continuous variable score for total initiating behavior requesting bids.

*Mullen Scales of Early Learning - AGS Edition* (Mullen, 1997). The Mullen is a measure that assesses performance across several areas of infant and early child development. Mullen subscales include: Gross Motor, Fine Motor, Visual Reception, Expressive Language, and Receptive Language. Each subscale yields a t score (0-50). The language subscales plus Fine Motor and Visual Reception can be combined into one comprehensive t score representing cognitive ability, which can then be converted to a standardized composite score, the Early Learning Composite \(M=100, SD=15\). The measure also provides age-equivalent scores, percentile ranks, and descriptive categories (i.e., Below Average, Above Average, etc.) relative to boys or girls within the child’s age group for each subscale and the Early Learning Composite score. Each of the four cognitive subscales assesses abilities from birth through 68 months. The Gross Motor subscale covers birth through 33 months. Items within each scale allow the examiner to assess strengths and weaknesses within domains. The Mullen has been shown to have solid psychometric properties, with good internal consistency, interrater reliability, and predictive validity (Mullen, 1997). Scoring occurs by the examiner during administration, but the task is videotaped so that the examiner can focus on the interaction and maintaining behavioral
momentum, and can then check scoring later if needed.

*Demographic Information Form.* The Demographic Information Form is a two-page questionnaire that asks questions through check-boxes and in short-answer format. Questions cover general family demographics, such as parent occupation, education level, and ethnicity; and a variety of risk factors that may increase vulnerability of a child to developmental problems. Such factors include pregnancy and birth complications, possible prenatal teratogens (parental smoking, alcohol, drug use, prescribed medications), prenatal trauma, anything outstanding in the child’s medical history, and any developmental or genetic disorders within the family history. The Demographic Information Form will be used qualitatively. Demographic information on individual children may be examined for the type of prenatal care, medical abnormalities or complications, and unusual aspects of family history. There are no specific hypotheses regarding the relationship of the Demographic Form to the target and comparison samples. Instead, information from the Demographic Form may be used in an exploratory fashion to examine individual profiles if it appears that this information will be useful and could generate potential questions and hypotheses for future studies.

*Training of Examiners*

Three examiners were trained by the Principal Investigator to administer the ESCS. Each examiner was already trained in Mullen administration, as the two primary examiners were doctoral-level school psychology students, and the third routinely gave Mullen assessments as part of a post-graduate research position.

Examiners were instructed to read the ESCS manual, and several meetings were held in sequence during the initial training. At the first meeting, the goals, techniques, and
procedure of ESCS administration were provided, and the examiners were encouraged to ask questions. The PI conducted three ESCS assessments with training babies on different days while the examiners observed live and then discussed the sessions in meetings directly following the assessments. After the initial training assessment, each examiner also conducted part of the ESCS with the other training babies, while the PI observed live and provided constructive feedback following the sessions.

Each examiner then conducted assessments on her own. Initial recordings of each session were reviewed by the PI, and constructive feedback was provided, ranging from details regarding specific rules of administration, assessment style, and ideas for maintaining momentum and/or managing difficult situations (i.e., parents who consistently interrupted, children who cried or refused to sit in parents’ laps). Once each examiner conducted the assessments according to standardized administration guidelines outlined in the ESCS manual and had improved management of difficult situations, she was allowed to administer official sessions. Administration of ESCS sessions was monitored throughout the year by weekly question and answer feedback, which primarily occurred over email and sometimes by phone, and by the PI randomly selecting and watching recorded sessions and providing feedback. Following a difficult situation or assessment issue for which an examiner received feedback, the PI watched a taped session to assess whether the examiner had corrected the situation, after which the PI provided feedback to the examiner about her improvement.

Mullen, Car Task, and Surprise Tasks assessments were also recorded and checked with standard administration instructions.

Lab Assessments

Assessments were conducted at the Frank Porter Graham Child Development Center
and at the TEACCH (The Education of Autistic and Communication-Handicapped Children) Center. When families arrived, they were provided with the consent form, a brief description of the project, and the demographic information form. All assessments were videotaped. The nonverbal communication tasks were administered first, followed by the Mullen developmental assessment. Total administration time lasted approximately 60-75 minutes. The examiner then met with the parent for a 5-10 minute debriefing session where she explained the goals of the project and answered any questions.

Coding of Data

*Interrater reliability.* The PI for this project served as the primary coder of the videotaped assessments of the ESCS. A second doctoral student was recruited as a reliability coder. She was not associated with the project but previously had been trained in ESCS coding, for which she had achieved high interrater reliability.

Sessions were randomly assigned to the PI and the reliability coder. Twelve sessions, or 21%, were randomly selected to serve as reliability sessions, and thus were coded by both coders. Coding was done using the Noldus Observer Event Recorder System 5.0, a professional, computerized scoring system designed for research to collect, manage, and analyze observational data. Behaviors were entered into the Observer System and pre-programmed by the PI with different computer keys corresponding to different behaviors. The Observer provides frequency counts for each behavior, along with a time stamp for when in the session that behavior was coded. The time stamp offers a huge advantage for reliability coding, in that it allows coders to review and compare discrepancies easily. Two sessions initially were coded for comparison and consensus-building on difficult interpretative decisions. From these sessions, several consistent errors were noted and corrected.
errors fell into the category of non-adherence to a coding rule specified in the ESCS manual and were easily clarified. Some errors involved interpretation of a type of behavior or gesture, and specific rules were discussed and consensus achieved on these behaviors.

Following these meetings, the remaining 10 sessions were coded by each of the two coders. Cohen’s Kappa, a common method of assessing interrater agreement, was selected as the measure of reliability because it calculates the probability associated with the amount of actual, observed agreement. Codes were judged as being in agreement if they were of the same category (i.e., joint attention versus behavior requesting) and if they fell within the same 4-second window. This method of agreement was more stringent than calculating only agreement in total number of joint attention versus behavior requesting codes, and was selected to ensure that coding was correct not only in total frequency of each type of behavior, but also as to whether the behavior was coded as social or nonsocial. Because Kappa sets a more stringent criterion for judging agreement than either intraclass correlations or percent agreement, a value of .70 or above is acceptable when using Kappa (Bakeman & Gottman, 1989).

Cohen’s Kappa was calculated for initiating joint attention versus initiating behavior requesting, for responding joint attention, and for responding behavior requesting. Kappa values for initiating joint attention versus initiating behavior requesting ranged from .20 to .89 across sessions, with an average value of .73, indicating good agreement. The session that received the Kappa value of .20 had very few joint attention and behavior requesting behaviors, making a high Kappa much more difficult to achieve. On this session, raters mismatched on only three instances, but because the overall number of possible behaviors was small, the calculated value resulted in a low Kappa. Without this session, the average
Kappa value for joint attention and behavior requesting increased to .79. For Responding Joint Attention, Kappa values ranged from .75 to 1.0, with an average Kappa of .98. For Responding Behavior Requesting, Kappa values ranged from .56 to 1.0, with an average Kappa of .93. The values for both types of responding behaviors indicated excellent overall agreement between raters.

Coding. The PI coded the ESCS using the Noldus Observer Event Recorder System. For actual coding and storage of data, the Observer System offers many advantages over paper and pencil coding from a VHS tape, including recording the exact time a behavior is coded, keeping track of the behaviors in separate computer files saved by the Child ID, and allowing the coder to pause or back track the disk at any time to clarify a code. Fifty-seven assessments were coded. The PI was blind to the status of the children and did not have access to risk scores or risk status until the completion of coding. One early assessment was deemed invalid due to the examiner allowing the child to sit on the floor, resulting in the child being too distracted to follow instructions or focus on any single task. The remaining 56 sessions were judged to be valid administrations. The Observer System was also used to code the Car and Surprise Tasks. Due to difficulties in completing all assessments before losing the child’s attention, several children had missing data from these two tasks, which were often conducted at the end of the session. In total, the PI coded 36 Car Task and 37 Surprise Task administrations.
CHAPTER 3
RESULTS

Preliminary Analyses

Sample characteristics

Descriptive statistics. Fifty-six children received valid laboratory evaluations. Of these, 25 met criteria for the target group (FYI score range = 11.000-31.125, $M = 17.81$), and 31 for the comparison group (FYI score range = 0.000-10.250, $M = 4.89$). Children in the two sample groups differed primarily on their FYI scores and were equivalent across many other salient factors, including: age at day of testing, gestational age at time of birth, and birth weight.

Children in the target group ranged in age at the day of testing from 12 months, 3 weeks to 14 months, 0 weeks, with a median age of 13 months, 1 week. Children in the comparison group ranged in age from 12 months, 3 weeks to 13 months, 2 weeks, also with a median age of 13 months, 1 week. All children were born at full-term (37-40 weeks gestation). For the target group, birth weight ranged from 5.4 to 10.4 pounds, with an average birth weight of 7.7 pounds. Birth weights for the comparison group ranged from 6.0 to 9.7 pounds, with an average birth weight of 7.8 pounds. Independent-samples t-tests comparing the two groups on age at day of testing, gestational age at time of birth, and birth weight were not significant.

---

2 Scores based on the new scoring that was modified after the sample was recruited.
Children were matched on gender, race, maternal education, and birth order, although the realities of recruitment and parent commitment to follow-through created some small disparities between the two groups. Both groups had a larger percentage of males versus females, but the target group consisted of a greater percentage of males than the comparison group, with 17 males and 8 females (68% male versus 32% female), versus 14 males and 17 females (45% male versus 55% female) for the comparison group. However, a chi-square test comparing the two groups on gender distribution was non-significant ($\chi^2 (1, N = 54) = 2.136, p = .144$). Table 2 presents the distribution of race, maternal education, and birth order between the two sample groups. As can be seen, minority groups were underrepresented as a whole in this sample, but the two sample groups were closely matched with each other on race. Children were also well-matched on maternal education, although children who had mothers with college or graduate work were overrepresented in this sample relative to the general population due to the recruitment area for this study. Finally, groups were fairly equivalent on birth order, although the comparison group had a slightly higher percentage of first-born children. Fisher’s Exact chi-square tests revealed no significant differences between groups on child race, maternal education, or birth order.

**Power Analysis.** A power analysis for linear regression was conducted to determine the likelihood of finding statistically significant differences between the target and comparison groups in order to help guide the interpretation of the findings of the study. A standard criterion was set at 80% power (Cohen, 1977). The effect size was estimated to be in the moderately-high range, given that the research literature indicates that typically-developing 13-14 month-old-children should show strong joint attention and requesting skills (Mundy, 1995). It is therefore reasonable to expect that children with delayed or deviant
development should show a significant failure in nonverbal communication by this age if not
developing appropriately in the realm of communication, as the autism literature indicates is
true of children who later develop autism (Mundy, 1995). The effect is estimated to be only
in the moderately-high range, however, given that there are still several unknown factors
about early detection of autism, the heterogeneity of the disorder, and the low incidence of
autism in the general population. Given this study’s sample size of 56 subjects and an
estimated effect size of .4, \( (d = .4) \), the power to detect significant differences if they exist
between groups, using linear regression analyses (alpha = .05, two-tailed) is estimated at
87%.

Examination of Variables

Description of Variables Used in this Study. Multiple variables were examined in this
study. As described under Measures, several subscale scores were derived from the Mullen
and were used in group comparisons to describe the sample. In addition, the independent
variables used in this study included the two Risk Groups (the target and comparison group)
and the child’s Risk Score. The dependent variables used in analyses were derived from the
ESCS. First, scoring for the ESCS yielded two types of variables. Some ESCS tasks were
scored using frequency counts for the total number of behaviors the child exhibited during
the task. Other tasks (primarily tasks targeting responding to joint attention or responding to
behavior requesting) were scored by awarding points for how many trials the child got
correct. For example, one task addressing responding to joint attention provides the child
with three trials, and the child can receive up to three points for a correct response.

As described shortly, frequency variables were transformed to rate variables due to
consistent differences in time spent on the measure per each examiner. For example, the
“Rate of Initiating Joint Attention” was derived from a frequency variable (Frequency of Initiating Joint Attention) that was converted to the rate of initiating joint attention behaviors per 5-minute segments. Percent Total Initiating Joint Attention and Percent Total Initiating Behavior Requesting were also calculated from the frequency scores in order to have a sense of what portion of the child’s communication served joint attention versus requesting functions. Also as described shortly, these scores were each combined with the rate variables to form two composite scores, one representing an Initiating Joint Attention Composite and one representing an Initiating Behavior Requesting Composite. Finally, scores for the tasks in which the child could receive an absolute number of points for right or wrong responses were combined by awarding the child a “percent correct” score. Each child then received a Percent Correct Responding to Joint Attention score and a Percent Correct Responding to Behavior Requesting score. Finally, as described under Measures, the Surprise Task yielded a categorical score representing the child’s response to a joint attention bid. All variables except the Surprise Task score were continuous. Tables 3 and 4 list and describe the independent and dependent variables examined in this study.

Visual Analysis. Histograms and boxplots were used to examine the distributions of individual dependent variables. Bivariate relationships were examined using scatterplots of dependent variables against child risk score to examine outliers and general patterns in the data.

Two outliers were identified from the univariate and bivariate plots. The most extreme and consistent outlier had an unusual pattern of a high risk score with high rates of behaviors. Coding notes from this case indicated that the examiner had cut short the session due to an uncooperative child. The second outlier represented the longest time spent for all
ESCS administrations, due to the examiner taking an excessive amount of time with an extremely inhibited child. Although this child’s responses would be deemed valid under an ordinary administration time frame, the length of the session justified this case’s removal from the data. In other words, this child’s behaviors may have been misrepresented by the examiner giving this child a much larger time frame during the assessment than is deemed appropriate for ESCS administration (i.e., well over 25 minutes). Both of these outliers represented target children. Because these outliers were outside the typical pattern of the data, coding notes indicated that these two sessions were problematic, and they appeared as outliers consistently, these two cases were eliminated from subsequent analyses.

The distribution of each dependent variable was examined for normality. Variables were examined for the entire sample and also according to risk group. The distribution shapes were consistent for whole sample versus split-sample graphs. That is, when the whole sample was distributed normally on a particular variable, both sample groups were also distributed normally on this variable, and vice versa. All dependent variables were normally distributed, with the exception of three variables. Percent Correct Initiating Behavior Requesting had a flat, jagged distribution for the entire sample. When examined by group, the target group had a flat distribution, and the comparison group had a steep slope and skewed tail at the higher percentages, suggesting that a few comparison children had very high scores in Responding Behavior Requesting.

Both variables representing rates of behaviors communicated during the Car Task (Rate of Car Task Initiating Joint Attention and Rate of Car Task Initiating Behavior Requesting) showed a strong floor effect, also with distributions skewed to the right. This pattern emerged for these two variables both for the entire sample and for the separate
groups. This abnormal pattern of scores likely occurred because of the short time frame allowed for this task, resulting in several children showing no or few behaviors during this task and resulting in a floor effect of scores piling up at 0 and 1. Although the length of the Car Task was constrained by the demands of conducting several tasks within a manageable time frame for this age group, it appears that for this task, a longer time frame might have elicited more behaviors, and thus, more normally distributed data. Results based on analyses that incorporate any of the three variables with abnormal distributions will be interpreted in light of this information.

**Score Transformations.** The variables obtained from both the ESCS and the Car Task represented frequencies of initiating and responding joint attention, and initiating and responding behavior requesting. These frequency scores were transformed in two ways for these analyses. The first transformation was to convert frequency scores into rate and percentage variables. This decision was made due to the finding that the three examiners had significantly different administration times on the ESCS. A 1-way ANOVA revealed a significant overall difference among the three: \( F(2, 50) = 29.287, p = .00 \). Post-hoc contrasts revealed significant differences between examiners 1 and 2 \( (p < .05) \), examiners 2 and 3, \( (p < .01) \), and examiners 1 and 3 \( (p < .01) \). These differences could create a confounding effect of administration time, so rate and percentage variables were calculated to control for time spent in ESCS administration. The frequency scores on the Car Task were also transformed to rate variables to match the ESCS variables.

The second transformation was to create composite scores from two individual scores on the ESCS that measured initiating joint attention, and two scores that measured initiating behavior requesting, in order to include best estimates in the comparisons. First, correlations
were calculated for each pair of scores to assess whether it would be suitable to combine them into a composite. The two variables that measured initiating joint attention (Rate of Initiating Joint Attention and Percentage of Initiating Joint Attention) were found to be significantly correlated: $r(54) = .67, p = .00$. These variables were standardized and combined using their mean values to form a composite score for initiating joint attention, which was labeled the Initiating Joint Attention (IJA) Composite. The two variables that measured initiating behavior requesting (Rate of Initiating Behavior Requesting and Percentage of Initiating Behavior Requesting) were also significantly correlated: $r(54) = .66, p = .00$. These variables were standardized and combined using their mean values into a composite score for initiating behavior requesting, which was labeled the Initiating Behavior Requesting (IBR) Composite.

**Risk-Group Criterion**

The risk-group criterion on the FYI is exploratory given the early stage of development for this measure. Ongoing research by the FYI authors will eventually provide a basis for defining specific levels of risk. For the present study, the criterion applied to the risk-score cut-off was the 80th percentile (FYI score of 11.000), based on the shape of the distribution of the 1300 FYIs collected at the initiation of this study. That is, children with risk scores of less than 11 would be in the comparison group, and children with scores of 11 or greater would be in the target group. It is possible that this criterion is not stringent enough given the low incidence of autism in the general population, and given that the sensitivity and specificity of the measure have yet to be firmly established. Therefore, the shape of the risk-score distribution of this dataset was examined to assess the possibility of a more appropriate risk-group cut-off based on a clear division in the data.
Considerations in establishing a new risk-group criterion included: (a) using visual analysis of the distributions to help locate a natural division of groups; (b) having relatively even sample sizes for each of the two comparison groups; (c) increasing the stringency of the criteria for the target group, given that the prevalence of autism in the population is small and the sample for this study was small; and (d) considering using a 50th percentile criterion for the comparison sample, given that 50th percentile represents a standard average score in the general population.

Applying these criteria led to a realignment of the sample into three groups: a low-risk group of children whose risk scores fell at the 50th percentile or less (risk score $\leq 5.5$), a high-risk group of children whose risk scores fell at the 90th percentile or greater (risk score $\geq 14.2$), and a moderate-risk group for children between the 51st and 89th percentiles (risk scores $\geq 5.5$ and $\leq 14.2$). This selection of criteria was based on visual determination of where natural groupings of the lowest and highest scores occurred for this sample, according to analysis of scatterplots (see Figure 2 for an example of the delineation of the 50th and 90th percentiles on the scatterplot for Risk Score and IJA Composite Score). These particular divisions allowed for equivalent sample sizes for the two extreme groups, resulting in a low-risk group of 18 children and a high-risk group of 16 children. Most important, the division of three groups made possible a second and more stringent comparison of the low-risk versus high-risk groups, and a more thorough testing of the hypotheses of this study, in that hypotheses were tested using both the original risk-group criterion as well as a second more stringent criterion.
Core Analyses

Developmental Data

As a first step in understanding the characteristics of the sample for this study, independent-samples t-tests were run comparing the risk groups on a range of developmental abilities as measured by the Mullen Scales of Early Learning. Mean differences were compared first on groups defined by the original risk-group criteria, and then the more extreme risk-group criteria.

Based on the original risk-group criteria that placed a child with an FYI score at the 80th percentile or above in the target group, results revealed a significant group difference, with the target group performing at a significantly lower level than the comparison group, for Expressive Language: \( t(48) = 2.125, p = .039 \). Group differences, again with the target group having lower scores than the comparison group, approached significance for Receptive Language: \( t(48) = -1.946, p = .057 \) and for Gross Motor Skills: \( t(48) = 1.799, p = .078 \). The direction of differences on the Fine Motor and Visual Reception subscales were also in the expected direction, with the target group having lower scores than the comparison group on both subscales. These group differences suggest that children with risk scores on the FYI at the 80th percentile or above generally showed slower development across areas of Receptive and Expressive Language, Gross and Fine Motor Skills, and Visual Reception, relative to children with lower risk scores.

Based on the more extreme risk-group criteria, the group difference on Gross Motor Skills was found to approach significance: \( t(29) = 2.005, p = .054 \). No significant group differences were found for other subscales, although the direction of differences remained the
same, with the high-risk children (at the 90th percentile or above) having lower scores across all subscales relative to the low-risk children (at the 50th percentile or lower).

**Hypothesis 1**

*Children in the target sample will perform significantly lower on measures of nonverbal communication relative to the comparison sample. Specifically, children in the target group will show lower scores on both joint attention and requesting skills relative to comparison children.*

This hypothesis was tested using a series of linear regression analyses and a chi square test. Risk Group was used to explain variance in several different scores representing initiating and responding joint attention, and initiating and responding behavior requesting, drawn from the ESCS and the Car Task. Group differences on the categorical variable representing responding joint attention, derived from the Surprise Task, were also examined. Variables that represented initiating and responding joint attention were: (a) Initiating Joint Attention Composite (ESCS); (b) Rate of Initiating Joint Attention (ESCS); (c) Rate of Car Task Initiating Joint Attention (Car Task); (d) Percent Correct Responding Joint Attention (ESCS); and (e) Ability to Respond to Joint Attention (Surprise Task). Variables that represented initiating and responding behavior requesting were: (a) Initiating Behavior Requesting Composite (ESCS); (b) Rate of Initiating Behavior Requesting (ESCS); (c) Rate of Car Task Initiating Behavior Requesting (Car Task); and (d) Percent Correct Responding to Behavior Requesting (ESCS).

All variables except the Surprise Score were continuous and were entered as dependent variables in a series of linear regression models. Risk Group was dummy-coded and entered as an independent variable. Pearson correlations were calculated to assess the
presence of colinearity between the independent variable and each dependent variable. The first series of regression analyses was conducted using the original risk-group criterion, followed by a second series based on the more stringent low-risk versus high-risk group criteria.

Original risk-group criteria. An examination of the correlations between Risk Group and each dependent variable revealed no significant correlations, indicating that Risk Group would not explain significant variance in initiating and responding joint attention and behavior requesting skills in regression models based on the original risk-group criteria. Results from the first series of regression analyses show that, when using the original risk-group criteria, variance in these communicative behaviors could not be explained significantly by Risk Group, as no single model achieved statistical significance. Table 5 presents model parameters for each dependent variable tested.

A chi-square analysis was also run to test group differences on the variable Ability to Respond to Joint Attention, drawn from the Surprise Task. Differences across categories were not significantly different between groups ($\chi^2 (3, n = 37) = 1.367, p = .713$). See Table 6 for counts and cell percentages for each group.

Extreme risk-group criteria. Analyses addressing Hypothesis 1 were re-run using the extreme risk-group criterion, comparing low-risk (50th percentile and below) to high-risk (90th percentile and above) infants. Pearson correlations indicated that one variable, Rate of Initiating Joint Attention, approached significance ($p = .095$) when correlated with Risk Group. Regression analyses using the extreme risk-groups revealed that more variance was accounted for in four variables when using the more stringent Risk Group criteria: Initiating Joint Attention Composite, Rate of Initiating Joint Attention, Percent Correct Responding to
Joint Attention, and Rate of Total Initiating Behaviors, although the models did not attain statistical significance. Table 7 presents these results. A chi-square test was run again comparing the Surprise Score between groups using the extreme risk-group criterion. Differences between groups remained non-significant: \( \chi^2 (3, n = 23) = 2.867, p = .413 \). Table 8 presents results from the chi-square analysis.

**Hypothesis 2**

Children in the target sample will show a deviant pattern of nonverbal communication, with stronger abilities in nonverbal requesting behaviors and weaker joint attention skills, relative to the comparison sample, which will show similar frequencies of the two types of behaviors.

Paired sample t-tests were used to address Hypothesis 2. First, statistical significance was tested within pairs of corresponding, joint attention/behavior requesting variables. Second, the significance levels of pairs for the target group were compared to the significance levels of pairs for the comparison group. Because it was believed that the target group would have splintered nonverbal communication abilities, it was hypothesized that the target group would show a significant difference within each variable pair (joint attention being significantly lower than behavior requesting). On the other hand, because it was believed that the comparison group would have relatively equivalent abilities in each type of nonverbal communication, it was hypothesized that variables within the comparison group pairs would fail to show significant differences from one another.

Variable pairs were drawn from the ESCS and the Car Task, the pairs representing Initiating Joint Attention with Initiating Behavior Requesting skills, and Responding Joint Attention with Responding Behavior Requesting skills. Specifically, three variable pairs were
tested: (a) Initiating Joint Attention Composite-Initiating Behavior Requesting Composite; (b) Rate of Car Task Initiating Joint Attention-Rate of Car Task Initiating Behavior Requesting; and (c) Percent Correct Responding Joint Attention-Percent Correct Responding Behavior Requesting.

This hypothesis addressed ideas that first, the two groups would show differing patterns of skills development, and second, the target group would show a pattern of weaker joint attention skills relative to behavior requesting.

**Original risk-group criteria.** Using the original risk-group criteria, results indicated that there was a significant difference for the comparison group, with Percent Correct Responding to Joint Attention being higher than Percent Correct Responding to Behavior Requesting, $t(27) = 3.064, p < .01$. For the target group, the difference between the Composite scores approached significance ($p = .092$), with the Initiating Joint Attention Composite being lower than the Initiating Behavior Requesting Composite. These results are presented in Table 9.

**Extreme risk-group criteria.** Hypothesis 2 analyses were re-run using the extreme risk groups, comparing pairs within the low-risk versus the high-risk groups. A significant difference between pairs emerged for the comparison group, with Percent Correct Responding to Joint Attention being significantly higher than Percent Correct Responding to Behavior Requesting, $t(16) = 3.193, p < .01$. A significant difference was also found for the target group, with the Initiating Joint Attention Composite being significantly lower than the Initiating Behavior Requesting Composite score, $t(15) = -2.471, p < .05$. Table 10 presents these results.

Based on the extreme risk-group criteria, joint attention, represented by the Initiating
Joint Attention Composite score, was significantly lower than behavior requesting, represented by the Initiating Behavior Requesting Composite score, for the high-risk group, whereas these two skills were not different for the low-risk group. This pattern of differences indicates that children with high-risk scores (90th percentile) show a different relationship between joint attention and behavior requesting skills relative to children with low-risk scores (50th percentile). Furthermore, this finding is in the predicted direction, indicating that children with high-risk scores show lower joint attention than behavior requesting skills, whereas children with low-risk scores show relatively equal skills in these two areas.

The other finding was that children with low-risk scores also showed a different pattern of skills for responding behaviors when compared to children with high-risk scores. Children in the low-risk group showed significantly stronger skills in Responding to Joint Attention relative to Responding to Behavior Requesting, whereas children in the high-risk group did not have a statistically significant difference between these two skills. Although this finding is not the same as the hypothesis suggests, it nevertheless supports the same pattern of strength differences between the two groups, in that responding joint attention was the stronger skill for the low-risk children, whereas the high-risk children did not show a strength in responding joint attention relative to behavior requesting. This result was found using either risk-group criteria.

**Additional Analyses**

**Identification of Examiner Effect**

The next step in examining the data was to compare scores obtained by the three examiners to explore the possibility that stylistic differences during testing had an effect on child behaviors. As noted earlier, frequency variables were converted to rate variables to
control for differences in time spent by examiners during ESCS administration. The identification of a potential effect above and beyond time was conducted first by examining scatterplots of three rate variables (Rate of Initiating Joint Attention, Rate of Initiating Behavior Requesting, Rate of Total Initiating Behavior), each plotted against the child’s Risk Score to compare examiners.

A notable difference was observed between the two primary examiners on the scatterplots. Rates of initiating behaviors obtained by Examiner 1 showed a negative linear relationship with risk score. For Examiner 2, however, the elicited behaviors were flat regardless of risk score, and for the variable Rate of Total Initiating Behavior, Examiner 2’s scores were positive and linear. Figure 3 presents this scatterplot (the most extreme example) with least-squares regression lines drawn for each Examiner’s cases to illustrate this visual difference. This visual distinction and intersection of least-squares regression lines justified testing for an interaction between examiner and rates of behaviors obtained for each risk group.

The seven cases tested by Examiner 3 did not follow a linear relationship but instead were more scattered. All seven cases had risk scores of < 15.000, and fell within the broad pattern for all cases with Risk Scores at 15.000 or less. At the onset of the study, this examiner intended to conduct an equivalent number of sessions. However, due to personal obligations, this person withdrew her participation. Therefore, the cases she tested were far fewer and without the variance that the other two examiners tested. However, she was included in the analyses.

Testing for an Interaction

Linear regression analyses were conducted to test for an interaction between
Examiner and Risk Group, with the rates of initiating behaviors as dependent variables. In the models, the independent variables included Examiner and Risk Group as main effects and Examiner × Risk Group as an interaction, all of which were dummy-coded for the regression analyses. Separate analyses were run for Rate of Initiating Joint Attention, Rate of Initiating Behavior Requesting, and Rate of Total Initiating Behaviors as dependent variables.

The analyses revealed a significant interaction between Examiners 1 and 2 for Rate of Initiating Joint Attention ($p < .05$) and Rate of Total Initiating Behaviors ($p < .05$). For all three measures of initiating behaviors, Examiner 1 elicited fewer behaviors from the target group than from the comparison group, whereas Examiner 2 elicited more behaviors from the target group relative to the comparison group. This effect also suggested that Examiner 1 generally elicited fewer behaviors than Examiner 2 for the target sample, and more behaviors than Examiner 2 for the comparison sample. These results clearly corresponded to the negative linear relationship seen for Examiner 1 in the scatterplots, compared with the relatively flat-line, and for one variable, a positive linear relationship, observed for Examiner 2. The interaction of Examiner by Group is illustrated in Figure 4, where the effect is seen most strongly for the Rate of Total Initiating Behaviors.

A significant interaction was also found between Examiners 1 and 3 for one variable, Rate of Total Initiating Behaviors ($p < .05$). Here, Examiner 1 elicited significantly fewer total initiating behaviors for the target group relative to the comparison group, whereas Examiner 3 elicited more behaviors from the target group relative to the comparison group. This interaction also meant that for total initiating behaviors, Examiner 1 generally elicited fewer behaviors than Examiner 3 for the target group but relatively more behaviors than Examiner 3 for the comparison group.
Analyses for Examiner Subsets

The finding of significant interactions between Examiners 1 and 2 and also a significant interaction for one variable between Examiners 1 and 3 justified examining data for each Examiner separately. Results indicated that there were no significant results for either hypothesis when only data from Examiner 2 or 3 were examined alone. However, significant linear relationships between variables were found for the subset of data tested by Examiner 1. These results for Examiner 1 are presented below. The relevance of the Examiner by Group interaction, and the meaning of the significant results for Examiner 1 alone is discussed in the final section of this paper.

Hypothesis 1. Examiner 1 tested a total of 29 cases, which included 10 children in the target group and 19 in the comparison group. A correlation matrix indicated that for Examiner 1’s cases, Risk Score was significantly correlated with the Rate of Initiating Joint Attention and the Rate of Total Initiating Behaviors, and correlated with the Initiating Joint Attention Composite at a level that approached statistical significance. The significant correlations indicate that Risk Score would explain significant variance in these variables.

A series of simple linear regressions were run to address hypothesis 1, based on the original risk-group criterion. The dummy-coded variable for Risk Group was entered as the independent variable, and several variables representing initiating and responding joint attention and requesting behaviors were entered as single dependent variables. Results indicated that Risk Group significantly predicted variance in the Rate of Initiating Joint Attention: \( t(26) = -2.271, p = .032 \), and the Rate of Total Initiating Behaviors: \( t(26) = -2.190, p = .038 \), and approached significance in explaining variance in the Initiating Joint Attention Composite score: \( t(26) = -1.937, p = .064 \). Table 11 presents these results.
Chi-square tests including only Examiner 1’s cases revealed non-significant
differences between the two groups across the categories of scores on the Surprise Task.

The regression and chi square analyses were re-run, using the extreme low versus
high-risk group criteria. This subsample included only 8 children in the high-risk group
versus 10 in the low-risk group. When using the extreme risk groups to explain variance in
the dependent variables, results were similar to the original risk groups, but with the
magnitude of differences increasing across all comparisons. Risk Group now explained
significant variance in Initiating Joint Attention Composite: $t(16) = -2.518, p = .023$), Rate of
Initiating Joint Attention: $t(16) = -2.705, p = .016$, and Rate of Total Initiating Behaviors:
$t(16) = -2.335, p = .033$, with the addition of one variable, Rate of Car Task Initiating Joint
Attention, approaching significant differences between groups: $t(12) = -1.885, p = .084$.
Model parameters for these and the non-significant variables are presented in Table 12. A
chi-square test conducted to test for group differences across categories of scores on the
Surprise Task using the extreme risk-group criteria did not result in any changes in the
findings, as results were still non-significant.

Overall, these results indicated that Risk Group did explain significant variance in
joint attention and behavior requesting variables for the subset of children tested by
Examiner 1. For these children, joint attention was found to be significantly lower for the
target group relative to the comparison group as assessed by the Rate of Initiating Joint
Attention. Examining this subset using the extreme risk group criteria resulted in larger
within-groups differences observed, with joint attention, as measured by the Rate of Initiating
Joint Attention and the Initiating Joint Attention Composite score, being significantly lower
for the high-risk group than for the low-risk group. Based on the subset of children tested by
Examiner 1, the null hypothesis can be rejected. These results differ from those based on the entire sample in that significant differences in joint attention skills are found as predicted between risk groups using either risk group criteria for the children tested by Examiner 1.

**Hypothesis 2.** Paired-sample t-tests were run to address Hypothesis 2 using only the subset of cases tested by Examiner 1. Results indicated that a significant difference emerged for the comparison group on the variable pair for responding behaviors: Percent Correct Responding Joint Attention-Percent Correct Responding Behavior Requesting, with the comparison children performing significantly better on responding joint attention than responding behavior requesting. This significant difference was found for the original risk-groups: $t(15) = 2.210, p = .043$, and it approached significance for the extreme risk-group criteria: $t(8) = 2.256, p = .054$. No other variable pairs attained significance. These results mirrored results from the entire sample for the Responding variable pair. However, when using only the subset of cases tested by Examiner 1, the significant differences seen for the Initiating Joint Attention/Initiating Behavior Requesting Composite score pair were no longer present, indicating the matched pair differences within the subset of cases for Examiner 1 were not strong enough to be detected. These results for Examiner 1, based on the original and the extreme risk-group criteria, are presented in Tables 13 and 14, respectively.

*Examination of Data as Continuum*

Regression analyses were also run using the child’s Risk Score, a continuous variable, as the independent variable, testing the variance Risk Score explained in each of the dependent variables above, with the idea that nonverbal communication might be best understand as a continuum in relation to risk, as opposed to group differences based on a risk-group cut-off score. Only the subset of the data tested by Examiner 1 was examined.
Results indicated that Risk Score significantly explained variance in three variables:

- Initiating Joint Attention Composite Score: $t(26) = -2.156, p = .040$
- Rate of Initiating Joint Attention: $t(26) = -2.802, p = .009$
- Rate of Total Initiating Behaviors: $t(26) = -3.020, p = .006$

Rate of Car Task Initiating Behaviors approached significance: $t(19) = -1.738, p = .098$.

Significant relationships are presented in Figure 5 as scatterplots with least squares regression lines. These results suggest that the data can be modeled as a single line in relation to Risk Score. In other words, the relationship between the child’s risk score and their joint attention abilities may be conceptualized as a continuum, with higher risk scores predicting lower scores on joint attention variables, rather than as two groups of children in different risk categories having mean differences on joint attention, without relationships between risk score and joint attention being seen within groups.
CHAPTER 4
DISCUSSION

Summary of Findings

The purpose of this study was to examine nonverbal communication skills in a cohort of children whose parents had reported a range of infant behaviors that are consistent with a developmental pattern that may precede an eventual diagnosis of autism. Specifically, do children who receive high risk scores on a recently-developed measure, the First-Year Inventory (FYI), demonstrate both a delayed and deviant pattern of nonverbal communication abilities relative to children who receive average scores on the FYI? The results of this study indicated that the answer to both questions is ‘yes’. Significant differences in both delay and deviancy were found for a subset of children tested by Examiner 1, and significant differences in deviancy were found as well for the entire sample.

Based on retrospective studies that have identified potential early indicators of autism, and extrapolating from what is known from research with older children with autism, it was hypothesized that 1-year-olds identified as being at-risk for autism would show, first, global delays in nonverbal communication, and second, a similar pattern of delay as older children with autism, with joint attention being more delayed than behavior requesting. The measure on which the sample was recruited, the FYI, is still in an early stage of development. Although the measure yields a total risk score, the cut-off for judging children to be at-risk for a later autism diagnosis has yet to be established conclusively. At the outset of this study,
based on the distribution of the 1300 FYIs collected at that time, the 80th percentile was selected as the best criterion to separate “at-risk” from “typical”. Because the 80th percentile is not a firmly established criterion, the distributions of risk scores against nonverbal communication variables of children in this sample were also examined to determine whether a more stringent cut-off might be more appropriate. This analysis suggested that comparing the 50th percentile of children with the 90th percentile might provide a more appropriate comparison of the two groups. Thus, the hypotheses for this study were addressed using both the original risk-group criterion of the 80th percentile, and the more extreme risk-group criteria comparing children with risk scores that fell at or below the 50th percentile with those whose scores fell at or above the 90th percentile.

Based on the original risk-group criterion for this study and looking at the entire sample, significant differences were not found on joint attention and behavior requesting skills between the target and comparison groups of children. A comparison of the more extreme groups (50th versus 90th percentiles) resulted in a greater magnitude of differences for some comparisons, with the high-risk group having lower scores than the low-risk group, although no single comparison reached the level of statistical significance. Therefore, when looking at the entire sample, statistically significant differences were not found between groups on overall levels of joint attention and behavior requesting.

Next, the second question was addressed for the entire sample. The relationship of joint attention and behavior requesting was examined within groups. Based on the original risk-group criteria, the Initiating Joint Attention Composite was lower than the Initiating Behavior Requesting Composite for the target group, although the difference only approached significance, relative to the comparison group, which showed a very similar level
of these two skills. The analysis also revealed a significant difference for the comparison
group for responding skills, with Responding Joint Attention being significantly higher than
Responding Behavior Requesting, whereas for the target group, responding skills were
relatively similar. When these comparisons were calculated using the more extreme risk-
group criteria, the composite score difference between skills was magnified for the target
group, rising to a level of significance, whereas the difference between composite scores for
the comparison group remained non-significant.

For Hypothesis 2, using either risk-group criteria revealed group differences in the
relationships of joint attention to behavior requesting, although the differences became more
pronounced with the use of the extreme risk-group criteria. The differences between the Joint
Attention Composite and Behavior Requesting Composite were as predicted, with joint
attention as measured by the composite score being lower relative to behavior requesting
only for the target group. The differences found for Responding skills were not exactly as
predicted because the split-skill difference was found with the comparison group instead of
the target group. However, this difference nonetheless supports the same general conclusion,
with the children deemed not at-risk for autism showing a pattern of stronger joint attention
skills relative to behavior requesting. It is possible that the target group showed the poorer
joint attention skills relative to behavior requesting for initiating skills and not responding
skills, because responding skills develop first (Mundy, 1995), and therefore, joint attention
may not have been as delayed relative to behavior requesting for responding skills as it was
for initiating skills.

Another aspect of this study was that, based on a visual analysis of the rates of
nonverbal communication behaviors shown by the children, the examiners appeared to have
elicited different results. An Examiner by Risk Group interaction was tested, with the finding of a significant interaction between the two primary examiners, and a significant interaction on one variable between one of the primary examiners and the third examiner. The data gathered by each examiner was then analyzed separately. Results indicated no significant findings for either hypothesis for Examiners 2 and 3. However, the subset of data gathered by Examiner 1 was strongly linear, with several nonverbal communication variables showing large differences between risk groups in the predicted direction.

Based on Examiner 1’s data, the results addressing Hypothesis 1 revealed statistically significant differences between the target and comparison groups for Rate of Initiating Joint Attention and Rate of Total Initiating Behaviors, and a difference between the Initiating Joint Attention Composite score that approached significance. Using the more extreme risk-groups criteria increased all group differences, with three variables - Rate of Initiating Joint Attention, Rate of Total Initiating Behaviors, and Initiating Joint Attention Composite - attaining significance, and one variable - Rate of Car Task Initiating Joint Attention - approaching significance.

Based on Examiner 1’s data, the results addressing Hypothesis 2 were similar to results based on the entire dataset, with the comparison group showing stronger Responding Joint Attention relative to Responding Behavior Requesting skills. However, the significant difference between composite scores was not found when using only Examiner 1’s data, and basing analyses on the more extreme group criteria did not effectively change results, although results were in the same direction as when based on the entire sample.

Taken together, when looking at overall nonverbal communication differences between groups, the direction of differences between groups consistently was found as
predicted, with children with higher risk scores showing lower means than children with lower risk scores. However, these differences did not reach the level of statistical significance, regardless of whether children were compared at the 80th percentiles, or between the 50th versus 90th percentiles, when analyzing the entire sample. Again when analyzing the entire sample, children in different risk groups did show differences in the relationship of joint attention to behavior requesting, with the target group showing weaker joint attention relative to requesting skills, and the comparison group showing stronger responding joint attention than behavior requesting skills. These differences were found using either risk-group criteria, though they were stronger when using the more extreme risk-group cut-off.

When analyses were conducted with only Examiner 1’s data due to an examiner effect, statistically significant nonverbal communication differences between risk groups emerged for 3 variables, and these differences became even stronger when comparing the more extreme groups, with a 4th variable showing near significant group differences. Specifically, these differences pertained to joint attention skills rather than behavior requesting, which is a finding that is both supported by the research literature and important in its own right in terms of the potential for joint attention as an early marker for autism risk. Finally, the relationship between skills for both groups was similar for Examiner 1 as for the entire sample, though the difference between composite scores was not found for the target group. Likewise, the finding of pattern differences between groups is important also in being supported by the research literature and in understanding how the groups differ developmentally.

These findings pertaining to nonverbal communication, the impact of increasing the
stringency of the risk-groups criterion, and the effect of examiner all will be discussed in
turn. In addition, the characteristics of the target sample recruited for this study based on the
nonverbal communication and developmental results is discussed. Limitations of this study
are discussed along with future directions research should take in these areas. Finally, overall
conclusions that can be drawn from this study are outlined.

Nonverbal Communication

Heterogeneity of sample

There are several possible explanations for the initial lack of significant findings
between groups on measures of nonverbal communication when using the entire sample.
First, if nonverbal communication is a true distinguishing characteristic of 1-year-olds at-risk
for an eventual diagnosis of autism, the lack of group differences in nonverbal
communication could indicate that the groups are not homogeneous. The group at-risk for
autism may in fact include children who have some, but not all, early indicators for autism,
and therefore might not demonstrate strong delays in nonverbal communication. The 80th and
90th percentiles used for risk-group cut-offs in this study might not result in a high enough
level of specificity, meaning that both criteria could also include children who have only
some characteristics of autism, or who have other types of developmental delays.

In fact, the assumption that children with high FYI scores have behaviors that place
them at risk for an eventual diagnosis of autism is based on the research literature and current
work that has been conducted on the FYI in the validation process. However, because the
FYI is yet to be validated and normed, it is too early to conclude that the sample of children
recruited with high FYI scores are a diagnostically-pure group. It is possible that children
with high FYI scores include children with developmental disabilities, or children who have
early indicators of autism but who will not go on to develop a profile of autism more broadly. The retrospective validity study conducted recently supports this possibility that the group of children with FYI scores in the 90th percentile and above includes children later diagnosed with autism and also those later diagnosed with other developmental disabilities (Watson et al., 2007). Follow-up of children who were screened with the FYI will be necessary to determine conclusively the outcome of children with a range of risk scores on this measure.

*Children with regressive-type autism*

Another possibility for the lack of statistically significant differences between the groups based on the entire sample is that the target group might have included children who tend to show some subtle signs of autism from the earliest months of life, but whose development of autism might have been different or regressive, relative to children with who tend to show a more consistent pattern of early deficits. Maestro and colleagues (2006) report on children with regressive autism who may have some, but not all, early indicators of autism, and therefore might show a different early pattern of nonverbal communication skill development relative to other children who show the nonregressive pattern of autism. Maestro and colleagues retrospectively compared children with early-onset autism, children with regressive autism, and children with typical development using early home videotapes (Maestro et al., 2006). Children with early-onset autism showed early signs of deficits in joint attention, whereas children with regressive autism showed a more typical development in joint attention until their first birthday, after which they tended to lose skills in joint attention at the same rate at which they developed skills in nonsocial attention (requesting skills). These children classified as having regressive autism also tended to show an especially strong interest for non-social stimuli during their first year, but did not show a
strong lack of interest in social stimuli until after their first birthday, nor did they show other positive autism signs at this time.

The issue of children with regressive autism confounding samples was seen in Werner et al.’s 2000 retrospective study, in which the authors failed to find social differences between children with autism and children with typical development at 12 months of age. However, when children whose parents had later reported regressive autism were removed from the sample, response to name distinguished the two groups, although communication and repetitive behaviors still did not distinguish the groups. It is possible that children with regressive autism might be picked up by the FYI because of a variety of other behaviors besides nonverbal communication that indicate an early manifestation of autism. In fact, other studies have reported that early deficits, including social deficits, are present for children who show a regressive pattern of autism (e.g., Ozonoff, Williams, & Landa, 2005). However, a specific delay in joint attention might not be evident until later in development for these children. In other words, children with regressive-type autism could have received high risk scores and therefore been recruited in this sample; however, they might not have shown strong deficits in nonverbal communication, thereby reducing differences between groups.

*Developmental trajectory of nonverbal communication*

Landa and Garrett-Mayer (2006) report in their prospective study of children later diagnosed with autism (non-regressive type) that the largest changes in developmental slowing were seen from 14 to 24 months of age, when overall developmental delays became more salient and the developmental trajectory slowed the most for children later diagnosed with autism versus those with language delays or typical development. These authors even
suggest that there might be a timing mechanism in the development of children with autism that is not linear, but that in some way creates more or less of a slowing of development at different ages (Landa & Garrett-Mayer, 2006). It is possible that nonverbal communication delay follows a non-linear path, and that the delays become more prominent after 12-14 months. If so, the 12 and 13-month-old children in this study who received high risk scores on the FYI would be only beginning to show nonverbal communication delays, which might not become as obvious until a few months beyond the time of this study.

**Significant findings with subset of data**

Finally, the reason significant findings may not have been found for the entire sample could be the presence of an examiner effect, discussed below. Statistically significant differences were found on several nonverbal communication variables for the subset of data tested by Examiner 1 (using either risk-group criteria). Given that the reduced sample size would have decreased power to detect significant differences, it is important to investigate these significant findings from Examiner 1. Primarily, the findings from Examiner 1’s data support the hypothesis that there are significant nonverbal communication differences between children with higher risk scores on the FYI relative to children with average risk scores. Importantly, these differences mainly were found for joint attention variables and were not seen for behavior requesting skills. This difference supports the idea that it is the social areas of development, including social nonverbal communication, which are the most impaired in very young children who may eventually receive an autism diagnosis. This finding is quite important is highlighting the role of nonverbal communication, and particularly joint attention, as an early area that can be assessed to identify infants at-risk for autism.
For this subset of data, as for the entire sample, pattern differences were found between groups on nonverbal communication. This finding of differences also supports the role that nonverbal communication has in early development of children at-risk for autism, and particularly joint attention, which was found to be lower than behavior requesting for the at-risk group.

*Extreme Risk-Group Criteria*

Using the more extreme 90th percentile instead of the 80th percentile to identify the at-risk group led to some stronger differences between groups (particularly when using only data from Examiner 1). The implications are that children with the highest risk scores may show the strongest deficits in nonverbal communication. Interestingly, as the risk-group criteria became more stringent, group differences on developmental variables, including Expressive and Receptive Language, became less obvious. One possibility for this finding is that children with the highest risk scores (≥90th percentile) may have been children who showed the most autistic-like behaviors, or who were at greatest risk for autism, whereas the group with elevated risk scores but less than 90th percentile may have included a greater number of children with developmental disabilities, who showed more extreme differences on developmental variables but did not show as typical of nonverbal communication pattern of autism.

The validity study recently conducted on the FYI lends some credence to this idea. Overall, the study found that a risk criterion set at a moderately high level (90th percentile in the study) showed high sensitivity in identifying all of the study sample’s children who had diagnoses of autism. However, it also resulted in low specificity, in that it also identified most of the children with developmental disabilities. When the criterion was raised to the 98th
percentile, the specificity improved in that the FYI did not identify any children with developmental disabilities. However, the sensitivity decreased, in that even though all the children identified had autism, many children with autism had FYI scores below the 98\textsuperscript{th} percentile.

It is possible that the sample for this study has a similar pattern, in that when the risk group criteria was raised, children now considered “high-risk” showed the more extreme differences in nonverbal communication that are typical of older children with autism. This higher risk-group criterion might have missed some children with elevated, but lower risk scores who would eventually develop autism, and it might also have missed some children with developmental disabilities who would not develop autism.

A final possibility for the lack of differences in developmental variables, such as language, that was observed using the extreme risk-group criteria is one of reduced sample size, which could decrease power and reduce statistical significance of findings. A comparison of effect sizes instead of significance levels shows that for Expressive Language, the effect size remains relatively similar for the two groups even though the significance level drops when using the extreme risk-group criteria (original risk-groups: $B = -4.732$; extreme risk-groups: $B = -4.332$). The effect size does decrease, however, for Receptive Language along with the significance level.

\textit{Examiner Effect}

This study also found an examiner effect. Specifically, Examiner 1 elicited fewer behaviors from the target sample relative to the comparison sample, but Examiner 2 elicited more behaviors from the target sample relative to the comparison sample. This effect was also found to a lesser extent with Examiner 3. Examiner 2 will be discussed first.
Several factors point to the data from Examiner 2 as being problematic. First, although Examiner 2 worked within an acceptable time frame, it was observed after multiple codings that because all of her sessions were conducted within the shortest recommended time (around 15 minutes), these assessments might not have allowed enough time for children in this sample to demonstrate representative behaviors. Although the frequency scores were converted to rate variables to control for time spent in administration, it appeared from scatterplots that when more time was allowed for each child, differences between groups emerged more clearly. It is possible that with the passage of time, the rate of communication actually increased for comparison children, whereas the rate of communication for target children remained more or less the same. It is also strongly possible that the sample recruited for this study demonstrated more subtle nonverbal communication differences relative to children who are typically assessed with the ESCS, who might already be grouped into non-autistic and clearly autistic groups. In order for the ESCS to draw out communication differences for children with high FYI scores who comprise a wider range of children, more time on the ESCS might have been necessary to allow these differences to emerge. Thus, it is possible that even though the ESCS allows for a range of time spent in administration, more time spent might actually allow for differences in nonverbal communication behaviors to emerge more clearly for children with subtle nonverbal communication differences, such as those in this sample.

A second problem was observed during coding of Examiner 2’s sessions, although this issue could not be confirmed until after determining which children belonged to which risk group. It appeared that when Examiner 2 sensed that a child was performing well, she tended to move on to the next task more quickly. When a child did not perform typically, she
seemed to stay with the task longer. This approach was probably based on her experience as an examiner and her clinical acumen in sensing when children were able to complete tasks, but it may have resulted in fewer behaviors being evoked from comparison children, whom she moved through key tasks that could have evoked more behaviors, too quickly. Although she had received feedback on allowing enough time for each child to communicate, her testing style may have been a subtle adjustment based on her sense of when the child was doing well. This manner of testing could have affected the results, with the comparison children tested by Examiner 2 actually showing fewer behaviors than the target children. Again, with a longer time frame for each child, a true differentiation of nonverbal communication abilities might have become more apparent.

A third possible explanation for the examiner effect could be the ways in which different examiner personalities and styles interacted with child temperament and ability level. Examiner 1 had a more relaxed approach when interacting with the infants, whereas Examiner 2 was more energetic and efficient. Early in the training, each examiner had received feedback about the aspects in which these styles were particularly problematic. For example, Examiner 1’s relaxed style initially caused her to slow behavioral momentum during testing and lose the interest of the more active children. On the other hand, Examiner 2 had received feedback about not being too aggressive in her interactions, causing the more inhibited children to withdraw. Each Examiner responded well to the respective feedback and adjusted her administrative style. However, it became more apparent during coding that these differing styles continued to exert an influence on children in different ways, depending on the child’s ability level, tolerance for the tasks, and/or temperament.

It is notable that there was nothing unusual or problematic that was observed in the
sessions or interaction of Examiner 3 with the children. Although a significant interaction was found on one variable between this examiner and Examiner 1, a look at the children tested by Examiner 3 suggests that this interaction might have disappeared had this person tested more children. The scores of the seven children tested by Examiner 3 were scattered and did not seem to be strongly divergent from the overall group of children who had similar risk scores of < 15.000. Because this examiner dropped out of the study due to personal obligations and was not able to test more children, it seemed appropriate not to draw strong conclusions about potential patterns in her data.

The differences between the two primary examiners lead to two important conclusions about working with children with disabilities, and particularly with autism. The first conclusion is that to elicit enough variability in nonverbal communication with children who may differ on these skills in subtle ways, it is important to spend enough time to allow these differences to emerge. In the same way that it is important for trained clinicians to have enough time to observe these communication differences in a clinical setting at this young age, it is also important to allow enough time for these subtle differences to emerge in a controlled research testing situation. This conclusion may apply to the diagnostic realm as well, where it might be important to spend enough time with younger children to make careful observations about subtle communication deficits.

The second lesson that can be drawn from the examiner effect is the importance of the human element of style and interpersonal interaction when testing very young children with or without autism or developmental disabilities. It became clearer through this project that factors such as demeanor, enthusiasm, pace, and energy level all had to be modulated according to the temperament and ability level of the child. It was noted from watching the
recorded sessions that some children responded better to a calm and reserved approach and tended to “shut-down” to a more aggressive style, whereas other children responded best to a high level of energy or tended to lose attention if the pace became too slow. In regard to children with autism, it is also possible that in general a more outgoing approach might cause them to withdraw, and for a child who is delayed, an approach that is too fast might not give them enough time to respond. Primarily, it is important for an examiner to have the skill of flexibility in testing, so that she can modify her behavior and style as the interaction progresses according to clinical acumen regarding how the child is responding to her.

*Autism and the Broader Profile*

It is noteworthy that based on the original risk-group criteria (80th percentile), children in the target group scored lower across language and gross motor areas relative to children in the comparison group. The finding of language differences is not surprising, given the assumption that children at-risk for developing autism would present in part with both receptive and expressive language delay. In fact, language is also one of the first delays often reported by parents of children with autism (Filipek et al., 1999). The lack of group differences found for Fine Motor and Visual Reception may indicate that fine motor skills deficits are not fully manifested by 13-14 months of age. Also, visually-based cognitive skills have even been found to be a strength in some autism samples (Schopler, 1993; Courchesne, Lincoln, Kilman, & Galambos, 1985). The gross motor skills group difference that was found is more surprising, given that children with autism have traditionally been thought of as not manifesting severe gross motor deficits. In fact, gross motor skills have been thought of as a way to differentiate children with autism from those with other types of developmental delays. There are several possible interpretations for the gross motor differences between
groups.

One possibility is that the gross motor differences emerged because the group with higher risk scores is heterogeneous and includes children not only who may receive an autism diagnosis but also children with developmental disabilities, including motor delays. This possibility is supported by the FYI-Retrospective study, which found that higher FYI scores did detect children with developmental disabilities in addition to children with autism diagnoses (Watson et al., 2007). A high level of sensitivity along with a lower level of specificity has been found with other early detection measures (e.g., M-CHAT), and is due to the difficulties of differential diagnosis at this early age, the fact that some early characteristics of autism are non-specific (e.g., language delay), and the fact that some known symptoms of autism, as discussed in the introduction, do not appear in typical development until the child is older (i.e., lack of pretend play, unusual use of language, etc.)

A second and perhaps more interesting possibility for the finding of gross motor differences is that children who present with early indicators of autism may in fact have gross motor skills delays as well. Recent literature that has explored this topic supports this idea. As previously discussed, Landa and Garrett-Mayer (2006) recently followed infants who had older siblings with autism prospectively using the Mullen, testing the infants at 6, 14, and 24 months on all subscales. At 24 months of age, the infants were tested with language measures and the Autism Diagnostic Observation Scale (ADOS) and classified as having autism, no autism, or having language delay. No significant subscale differences were found at the 6 month testing, but at 14 months, children who were later classified within the autism group had significantly lower scores relative to children with typical development on all subscales (including Gross Motor), except Visual Reception. Furthermore, by 24 months, the ASD
group performed significantly worse on all domains, and also could be distinguished at this time from the language-delayed group on significantly lower Gross Motor, Fine Motor, and Receptive Language scores. Landa and Garrett-Mayer suggest that early gross motor delays, an area of development not traditionally thought of as being linked to autism, may be an important early deficit.

Other authors have found various types of motor delay in older children with autism, such as clumsiness (Ghaziuddin & Butler, 1998), movement abnormalities (Eisenmajor et al., 1998), motor control (Jansiewicz et al., 2006), and motor dysfunction such as problems with proprioception, motor planning, praxis and body mapping, and reaction time (Hill & Leary, 1993; Minshew, Goldstein, & Seigel, 1997; Rogers, Bennetto, McEvoy, & Pennington, 1996). It is possible that gross motor deficits in fact also could be significant early precursors or indicators of children who later receive an autism diagnosis.

Recent research has been conducted on this topic. Teitelbaum (2002) has explored the possibility of using an early gross motor skills assessment to identify infants at-risk for autism. Teitelbaum analyzed movements from retrospective home videotapes in infants less than one year of age who were later diagnosed with autism. Compared to children with typical development, children later diagnosed with autism had disrupted development of certain reflexes, with some reflexes continuing beyond expected termination, such as the asymmetrical tonic neck reflex being observed at eight months of age. In contrast, other reflexes did not appear when they should have, including protective reflexes (holding out hands to catch oneself when falling) not being present at eight months, and the head-verticalization reflex (head and neck adjusting vertically when one’s body tilts to one side) not being present by 6-8 months of age. Because these children were compared to typically-
developing children, the results of this study may not be specific to autism but may be
generalized to children with other types of developmental disabilities as well. Nevertheless,
Teitelbaum speculates that it is possible that autism spectrum disorders could be detected as
early as 6 months if disrupted reflexes are identified and analyzed at this time.

Limitations

Identification of representative sample

As discussed earlier, minority and lower socioeconomic groups were quite
underrepresented in this sample. This was due to several factors, including the demographics
of families who tended to respond to the FYI mailing packet and the demographics of those
living in the immediate area. Although it is not currently believed that autism is present to
any greater or lesser extent across various ethnic or socioeconomic groups, cultural factors
nevertheless could have strong effects on the identification, assessment, and treatment of
autism, and therefore represent a limitation of this study. In particular, how parents interpret
items on the FYI are likely influenced by cultural factors and represent an area that this study
could not address. Furthermore, issues such as attitudes about reporting developmental
problems in one’s child, level of knowledge about early development, cultural differences in
interpretation of early symptoms and child behaviors, and financial and other resources
required to seek out and follow-through with research studies, and more important, with
assessment and treatment, are some of the factors that may influence diagnostic and
treatment practices for children with autism.

Number of Measures

A final limitation of this study was the ambitious goal of using several measures
within a manageable time frame for 1-year-olds. Examiners found it challenging to maintain
children’s attention and interest in tasks beyond about a one-hour window. Therefore, tasks that were administered last were often rushed or did not yield enough data to be used to their fullest extent. The primary example was the Car Task, which showed a strong floor effect for both the target and comparison samples, with many children not showing any communication behaviors regardless of group membership. This task was often rushed at the end of the assessment due to children losing interest, even though the task in itself would typically evoke interest in a 1-year-old.

**Future Directions**

Future research needs to focus on recruitment of samples of 1-year-olds who are at-risk for receiving an autism diagnosis using early screening measures such as the FYI. Future work will focus on continuing the process of validation of the FYI. Follow-up research is needed for children identified as being at-risk for autism based on their FYI score, as well as for children identified as not being at-risk according to the FYI, to further establish developmental pathways of autism and typical development. Research should also continue to investigate what early behaviors discriminate children at-risk for autism versus those with developmental disabilities (i.e., non-response to name, discussed in the introduction). Such research will further refine the FYI to achieve high sensitivity and specificity. With this goal achieved, the FYI could be used to recruit a more homogeneous sample with which testing could be conducted on nonverbal communication and other important early deficits associated with a later diagnosis of autism.

Sampling should focus on recruitment of representative samples along with continued efforts to ensure follow-through with appointments. An upcoming Spanish translation of the FYI will greatly aid the efforts to reach the Spanish-speaking population. Other efforts could
focus on additional incentives for underrepresented groups to return the FYI packets by mail to aid in the collection of data from a broader cross-section of cultural and socioeconomic groups. Offering support such as transportation or Saturday sessions might enable some families to participate more easily in lab-based assessments.

Measurement of deficit areas such as nonverbal communication needs to proceed with a planful approach, with the tolerance of the child taken into account and balanced with spending enough time on key aspects of the assessment, so that the time spent yields the greatest amount of information. Examiners need to practice flexibility so that they pace themselves according to the child’s ability level. They should exercise good clinical judgment regarding the child’s responses to their tasks, and, importantly, to the examiner’s personality and style of interaction, adjusting their energy level, pace, or approach according to how the child is responding to them.

Follow-up assessment needs to include measures of nonverbal communication to better understand this early developmental and foundational skill, which could be a strong and easily identifiable early indicator of autism. Longitudinal research on the development of nonverbal communication in children at-risk for autism would enable researchers to better understand the emergence of this deficit and the timing at which nonverbal communication develops in children with autism spectrum disorders.

A final area that needs to be examined is the longitudinal progression of other developmental domains for infants identified as being at-risk for autism, particularly gross motor skills. The developmental relationships between gross motor skills and other areas believed to be problematic in the early development of autism could provide valuable information for researchers, clinicians, and families about early identification of autism.
Conclusion

This study provided a unique opportunity to identify 1-year-olds whose parents reported a constellation of behaviors that are believed to precede a likely eventual diagnosis of autism. Thus far, it has not been possible to study nonverbal communication in a sample of children this young who have early indicators of autism, because screening measures are generally not available for this age group. This makes the communication profiles of such children difficult to study in detail, with only broad generalizations possible about their early nonverbal communication skill patterns.

A finding of nonverbal communication delay could assist in identifying children who are at-risk for autism, in that these skills are measurable at this age group. Furthermore, a pattern of nonverbal communication deviance might also help distinguish children at-risk for autism from children at-risk for other developmental delays. This study has provided evidence of both nonverbal communication skills delays and deviance in a group of children who were younger at age of testing than most research samples.

The findings of group differences in nonverbal communication were statistically significant for a subset of children tested by one of the primary examiners, and they support the hypothesis that there are nonverbal communication differences between children with high risk scores on the FYI relative to children with average scores. Importantly, significant differences were found for joint attention variables as opposed to behavior requesting variables, which provides additional evidence that social areas of nonverbal communication are in fact delayed in young children who may receive an eventual diagnosis of autism. In addition, the finding of differences in the pattern of nonverbal communication skills for the overall sample further strengthens the idea that the sample groups were different on this key
variable. Furthermore, the pattern differences were as predicted, with joint attention being weaker than behavior requesting for the target group but not for the comparison group.

The findings of this study on the overall sample are tempered by the likelihood that the target sample may have included some children with other developmental delays besides autism, or children falling along the autism spectrum who did not actually rise to a level that would warrant an eventual autism diagnosis. In addition, the influence of examiner style and time spent during testing seemed to influence the data. Nevertheless, the findings from this study support the notion that children with at-risk scores on the FYI differ along nonverbal communication dimensions.

This study also found that the risk groups on the FYI (separated by the 80th percentile) were statistically different on developmental skills, including Expressive Language, Receptive Language, and Gross Motor Skills. These developmental differences indicate that the two groups identified and discerned from the parent-reported FYI risk score are in fact different developmentally when tested in a laboratory context, and that the profile of differences includes language and motor deficits seen in the risk group. Given that the children identified by the FYI have behaviors that are believed to precede an autism diagnosis, it is important to try to characterize them on the basis of additional developmental variables as well. Alternatively, this developmental profile can be used to try to characterize the children who are being flagged by the FYI as being “at-risk”.

Future work on identification of infants at-risk for autism and on a better understanding of the earliest deficits and developmental profiles can aid in providing information, support, and early intervention for these families as well as increasing our conceptualization and knowledge of this developmental disorder.
Table 1
List of joint attention and behavior requesting nonverbal communication gestures coded during the Early Social Communication Scales (ESCS) assessment

<table>
<thead>
<tr>
<th>Joint Attention</th>
<th>Behavior Requesting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initiating Joint Attention</td>
<td>Initiating Behavior Requesting</td>
</tr>
<tr>
<td>Joint Attention Eye Contact</td>
<td>Behavior Requesting Eye Contact</td>
</tr>
<tr>
<td>Joint Attention Alternates Eye Contact</td>
<td>Behavior Requesting Reach</td>
</tr>
<tr>
<td>Joint Attention Point</td>
<td>Behavior Requesting Reach + EC</td>
</tr>
<tr>
<td>Joint Attention Point + Eye Contact</td>
<td>Behavior Requesting Point</td>
</tr>
<tr>
<td>Joint Attention Show</td>
<td>Behavior Requesting Point + Eye Contact</td>
</tr>
<tr>
<td></td>
<td>Behavior Requesting Give</td>
</tr>
<tr>
<td></td>
<td>Behavior Requesting Give + Eye Contact</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Responding Joint Attention</th>
<th>Responding Behavior Requesting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Look trials (follows distal point: side)</td>
<td>Follows command: Give with gesture</td>
</tr>
<tr>
<td>Look trials (follows distal point: back)</td>
<td>Follows command: Give without gesture</td>
</tr>
<tr>
<td>Book trials (follows proximal point)</td>
<td></td>
</tr>
</tbody>
</table>
Table 2
Demographics of sample for Child Race, Maternal Education, and Birth Order

<table>
<thead>
<tr>
<th>Child Race</th>
<th>Caucasian</th>
<th>African-Am.</th>
<th>Asian</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>Target</td>
<td>21</td>
<td>84%</td>
<td>1</td>
<td>4%</td>
</tr>
<tr>
<td>Comparison</td>
<td>26</td>
<td>84%</td>
<td>1</td>
<td>3%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Maternal Education</th>
<th>HS Education</th>
<th>Some College</th>
<th>Bachelors</th>
<th>Graduate</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>Target</td>
<td>0</td>
<td>0%</td>
<td>2</td>
<td>8%</td>
</tr>
<tr>
<td>Comparison</td>
<td>1</td>
<td>3%</td>
<td>3</td>
<td>10%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Birth Order</th>
<th>First-born</th>
<th>Second-born</th>
<th>Third-born</th>
<th>Fourth-born</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>Target</td>
<td>13</td>
<td>52%</td>
<td>7</td>
<td>28%</td>
</tr>
<tr>
<td>Comparison</td>
<td>20</td>
<td>65%</td>
<td>7</td>
<td>3%</td>
</tr>
</tbody>
</table>
### Table 3

List of independent variables and descriptions

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original Risk Group</td>
<td>Categorical variable; defines two groups delineated by 80\textsuperscript{th} percentile. Target group = 80\textsuperscript{th} percentile or higher; comparison group = 79\textsuperscript{th} percentile or lower.</td>
</tr>
<tr>
<td>Extreme Risk Group</td>
<td>Categorical variable; defines more stringent criteria for two groups. High-risk group = 90\textsuperscript{th} percentile or higher; low-risk group = 50\textsuperscript{th} percentile or lower.</td>
</tr>
<tr>
<td>Risk Score</td>
<td>Continuous variable; child’s score obtained on the FYI; ranges from 0-50.</td>
</tr>
</tbody>
</table>
### Table 4

List of dependent variables and description

<table>
<thead>
<tr>
<th>Dependent Variables</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Initiating Joint Attention (IJA) Composite</td>
<td>1. Composite score created with Rate of Initiating Joint Attention and Percent Total Initiating Joint Attention.</td>
</tr>
<tr>
<td>2. Rate of Initiating Joint Attention</td>
<td>2. Rate of Initiating Joint Attention behaviors the child communicates per 5-minute segments.</td>
</tr>
<tr>
<td>3. Rate of Car Task Initiating Joint Attention</td>
<td>3. Rate of Joint Attention behaviors the child communicates per 1-minute segments during the Car Task.</td>
</tr>
<tr>
<td>4. Percent Correct Responding JA</td>
<td>4. Percent correct the child obtains from responding to joint attention bids, both proximal point and distal point.</td>
</tr>
<tr>
<td>5. Ability to Respond to JA (Surprise Task Score)</td>
<td>5. Categorical score indicating child’s response to JA bid: head turn, head turn plus affect, head turn, point, &amp; affect</td>
</tr>
<tr>
<td>6. Initiating Behavior Requesting Composite</td>
<td>6. Composite score created from Rate of Initiating Behavior Requesting and Percent Total Initiating Behavior Requests</td>
</tr>
<tr>
<td>7. Rate of Initiating Behavior Requesting</td>
<td>7. Rate of Initiating Behavior Requesting behaviors the child communicates per 5-minute segments.</td>
</tr>
<tr>
<td>8. Rate of Car Task Initiating Beh Requesting</td>
<td>8. Rate of Behavior Requesting the child communicates per 1-minute segments during the Car Task</td>
</tr>
<tr>
<td>9. Percent CorrectResponding Beh Req.</td>
<td>9. Percent correct the child obtains from responding to behavior requesting “Give it to me” commands</td>
</tr>
<tr>
<td>10. Rate of Total Initiating</td>
<td>10. Rate of Total Initiating (JA + BR) per 5 min. segments</td>
</tr>
</tbody>
</table>
Table 5

Linear regression models of explained variance in individual dependent variables by risk group; Original risk-group criteria

<table>
<thead>
<tr>
<th>Risk Group</th>
<th>IJAComp</th>
<th>RateJA</th>
<th>CarIJA</th>
<th>PctRJA</th>
<th>IBRComp</th>
<th>RateBR</th>
<th>CarIBR</th>
<th>PctRBR</th>
<th>RateTot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unstand. coeff</td>
<td>-.311</td>
<td>.776</td>
<td>.526</td>
<td>-2.243</td>
<td>.333</td>
<td>-.834</td>
<td>.189</td>
<td>10.943</td>
<td>-.039</td>
</tr>
<tr>
<td>Standard error</td>
<td>.234</td>
<td>.794</td>
<td>.678</td>
<td>6.001</td>
<td>.249</td>
<td>.853</td>
<td>.551</td>
<td>8.792</td>
<td>1.141</td>
</tr>
<tr>
<td>p value</td>
<td>.190</td>
<td>.333</td>
<td>.443</td>
<td>.710</td>
<td>.188</td>
<td>.333</td>
<td>.734</td>
<td>.219</td>
<td>.973</td>
</tr>
</tbody>
</table>

| Constant   |         |        |        |        |         |        |        |        |         |
| Unstand. coeff | .068    | 5.084  | 1.472  | 56.935 | -.110   | 8.035  | 1.535  | 37.689 | 13.089  |
| Standard error | .154    | 1.304  | .480   | 3.889  | .164    | 1.401  | .390   | 5.617  | 1.875   |
| p value     | .661    | .000** | .004** | .000** | .507    | .000** | .000** | .000** | .000**  |

Model $R^2$ | .033    | .018   | .017   | .003   | .034    | .018   | .003   | .032   | .000    |

* $p < .05$; **$p < .01$
Table 6

Chi-square for surprise score differences between groups; Original risk-group criteria

<table>
<thead>
<tr>
<th>Risk Group</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Count</td>
<td>7</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>15</td>
</tr>
<tr>
<td>% within group</td>
<td>46.7%</td>
<td>20.0%</td>
<td>13.3%</td>
<td>20.0%</td>
<td>100.0%</td>
</tr>
<tr>
<td>Comparison Group</td>
<td>7</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>22</td>
</tr>
<tr>
<td>Count</td>
<td>31.8%</td>
<td>18.2%</td>
<td>27.3%</td>
<td>22.7%</td>
<td>100.0%</td>
</tr>
<tr>
<td>Total</td>
<td>14</td>
<td>7</td>
<td>8</td>
<td>8</td>
<td>37</td>
</tr>
<tr>
<td>Count</td>
<td>37.8%</td>
<td>18.9%</td>
<td>21.6%</td>
<td>21.6%</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

\[ \chi^2 = 1.367, \ p = .713 \]
Table 7

Linear regression models of explained variance in individual dependent variables by risk group; Extreme risk-group criteria

<table>
<thead>
<tr>
<th>Risk Group</th>
<th>IJAComp</th>
<th>RateJA</th>
<th>CarIJA</th>
<th>PctRJA</th>
<th>IBRComp</th>
<th>RateBR</th>
<th>CarBR</th>
<th>PctRBR</th>
<th>RateTot</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unstand. coeff</td>
<td>-.496</td>
<td>1.836</td>
<td>.506</td>
<td>-7.622</td>
<td>.244</td>
<td>-.150</td>
<td>-.081</td>
<td>7.162</td>
</tr>
<tr>
<td></td>
<td>p value</td>
<td>.105</td>
<td>.095</td>
<td>.622</td>
<td>.284</td>
<td>.416</td>
<td>.891</td>
<td>.904</td>
<td>.558</td>
</tr>
<tr>
<td></td>
<td>Standard error</td>
<td>.204</td>
<td>1.719</td>
<td>.730</td>
<td>4.626</td>
<td>.204</td>
<td>1.751</td>
<td>.481</td>
<td>7.960</td>
</tr>
<tr>
<td></td>
<td>p value</td>
<td>.818</td>
<td>.062</td>
<td>.043*</td>
<td>.000**</td>
<td>.653</td>
<td>.000**</td>
<td>.002**</td>
<td>.000**</td>
</tr>
<tr>
<td>Model $R^2$</td>
<td>.080</td>
<td>.084</td>
<td>.012</td>
<td>.038</td>
<td>.021</td>
<td>.001</td>
<td>.027</td>
<td>.012</td>
<td>.036</td>
</tr>
</tbody>
</table>

* $p < .05$; ** $p < .01$
Table 8

Chi-square for surprise score differences between groups; Extreme risk-group criteria

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High Risk Group</strong></td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>Count</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>% within group</td>
<td>40.0%</td>
<td>30.0%</td>
<td>20.0%</td>
<td>10.0%</td>
<td>100.0%</td>
</tr>
<tr>
<td><strong>Low Risk Group</strong></td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td>Count</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td>% within group</td>
<td>23.1%</td>
<td>15.4%</td>
<td>23.1%</td>
<td>38.5%</td>
<td>100.0%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>7</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>23</td>
</tr>
<tr>
<td>Count</td>
<td>7</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>23</td>
</tr>
<tr>
<td>% within group</td>
<td>30.4%</td>
<td>21.7%</td>
<td>21.7%</td>
<td>26.1%</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

χ² = 2.867, p = .413
Table 9

Paired sample t-tests comparing pairs of initiating and responding variables; Original risk-group criteria

<table>
<thead>
<tr>
<th>Variable Pair</th>
<th>Paired Mean Diff</th>
<th>t score</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Target Group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IJA-IBR Composites</td>
<td>- .466</td>
<td>-1.759</td>
<td>22</td>
<td>.092</td>
</tr>
<tr>
<td>Rate Car IJA-IBR</td>
<td>.275</td>
<td>.394</td>
<td>17</td>
<td>.698</td>
</tr>
<tr>
<td>PctCorrRJA/RBR</td>
<td>8.205</td>
<td>1.113</td>
<td>18</td>
<td>.281</td>
</tr>
<tr>
<td><strong>Comparison Group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IJA-IBR Composite</td>
<td>.177</td>
<td>.529</td>
<td>29</td>
<td>.601</td>
</tr>
<tr>
<td>Rate Car IJA-IBR</td>
<td>-.062</td>
<td>-.136</td>
<td>17</td>
<td>.893</td>
</tr>
<tr>
<td>PctCorrRJA-RBR</td>
<td>18.011</td>
<td>3.064</td>
<td>27</td>
<td>.005**</td>
</tr>
</tbody>
</table>

*p < .05; **p < .01
**Table 10**

Paired sample t-tests comparing pairs of initiating and responding variables; Extreme risk-group criteria

<table>
<thead>
<tr>
<th>Variable Pair</th>
<th>Paired Mean Diff.</th>
<th>t score</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High-Risk Group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IJA-IBR Composites</td>
<td>-7.86</td>
<td>-2.471</td>
<td>15</td>
<td>.026*</td>
</tr>
<tr>
<td>Rate Car IJA-IBR</td>
<td>.4221</td>
<td>.429</td>
<td>11</td>
<td>.676</td>
</tr>
<tr>
<td>PctCorrRJA/RBR</td>
<td>13.600</td>
<td>1.586</td>
<td>11</td>
<td>.141</td>
</tr>
<tr>
<td><strong>Low-Risk Group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IJA-IBR Composite</td>
<td>-.045</td>
<td>-.106</td>
<td>17</td>
<td>.916</td>
</tr>
<tr>
<td>Rate Car IJA-IBR</td>
<td>-.165</td>
<td>-.241</td>
<td>10</td>
<td>.815</td>
</tr>
</tbody>
</table>

*p < .05; **p < .01
### Table 11

Examiner 1 cases only: Linear regression models of explained variance in individual dependent variables by risk group; Original risk-group criteria

<table>
<thead>
<tr>
<th>Risk Group</th>
<th>IJACComp</th>
<th>RateJA</th>
<th>CarIJA</th>
<th>PctRJA</th>
<th>IBRCcomp</th>
<th>RateBR</th>
<th>CarIBR</th>
<th>PctRBR</th>
<th>RateTot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unstand. coeff</td>
<td>-.617</td>
<td>-2.544</td>
<td>-.488</td>
<td>1.813</td>
<td>.099</td>
<td>-.827</td>
<td>-.067</td>
<td>15.002</td>
<td>-3.383</td>
</tr>
<tr>
<td>Standard error</td>
<td>.319</td>
<td>1.120</td>
<td>.464</td>
<td>8.757</td>
<td>.343</td>
<td>1.131</td>
<td>.716</td>
<td>12.900</td>
<td>1.544</td>
</tr>
<tr>
<td>p value</td>
<td>.064</td>
<td>.032*</td>
<td>.306</td>
<td>.838</td>
<td>.775</td>
<td>.471</td>
<td>.927</td>
<td>.257</td>
<td>.038*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Constant</th>
<th>IJACcomp</th>
<th>RateJA</th>
<th>CarIJA</th>
<th>PctRJA</th>
<th>IBRCcomp</th>
<th>RateBR</th>
<th>CarIBR</th>
<th>PctRBR</th>
<th>RateTot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unstand. coeff</td>
<td>.269</td>
<td>7.502</td>
<td>1.259</td>
<td>56.455</td>
<td>-.143</td>
<td>6.600</td>
<td>1.452</td>
<td>36.322</td>
<td>14.113</td>
</tr>
<tr>
<td>Standard error</td>
<td>.190</td>
<td>.669</td>
<td>.268</td>
<td>5.152</td>
<td>.205</td>
<td>.676</td>
<td>.413</td>
<td>7.298</td>
<td>.923</td>
</tr>
<tr>
<td>p value</td>
<td>.169</td>
<td>.000**</td>
<td>.000**</td>
<td>.000**</td>
<td>.492</td>
<td>.000**</td>
<td>.002**</td>
<td>.000**</td>
<td>.000**</td>
</tr>
</tbody>
</table>

Model $R^2$ | .126 | .166 | .055 | .002 | .003 | .020 | .000 | .056 | .156 |

* $p < .05$; ** $p < .01$
**Table 12**

Examiner 1 cases only: Linear regression models of explained variance in individual dependent variables by risk group; Extreme risk-group criteria

<table>
<thead>
<tr>
<th>Risk Group</th>
<th>IJAComp</th>
<th>RateJA</th>
<th>CarJJA</th>
<th>PctRJA</th>
<th>IBRCcomp</th>
<th>RateBR</th>
<th>CarIBR</th>
<th>PctRBR</th>
<th>RateTot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unstand. coeff</td>
<td>-.982</td>
<td>-3.968</td>
<td>-.918</td>
<td>-4.768</td>
<td>.238</td>
<td>-.904</td>
<td>-.183</td>
<td>11.339</td>
<td>-4.892</td>
</tr>
<tr>
<td>Standard error</td>
<td>.390</td>
<td>1.467</td>
<td>.487</td>
<td>8.258</td>
<td>.415</td>
<td>1.433</td>
<td>.952</td>
<td>19.239</td>
<td>2.096</td>
</tr>
<tr>
<td>p value</td>
<td>.023*</td>
<td>.016*</td>
<td>.084</td>
<td>.572</td>
<td>.575</td>
<td>.537</td>
<td>.850</td>
<td>.566</td>
<td>.033*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Risk Group</th>
<th>Constant</th>
<th>IJAComp</th>
<th>RateJA</th>
<th>CarJJA</th>
<th>PctRJA</th>
<th>IBRCcomp</th>
<th>RateBR</th>
<th>CarIBR</th>
<th>PctRBR</th>
<th>RateTot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unstand. coeff</td>
<td>.468</td>
<td>8.432</td>
<td>1.451</td>
<td>66.589</td>
<td>-.163</td>
<td>6.837</td>
<td>1.678</td>
<td>38.661</td>
<td>15.289</td>
<td></td>
</tr>
<tr>
<td>p value</td>
<td>.091</td>
<td>.000**</td>
<td>.001**</td>
<td>.000**</td>
<td>.563</td>
<td>.000**</td>
<td>.020*</td>
<td>.007**</td>
<td>.000**</td>
<td></td>
</tr>
</tbody>
</table>

| Model R² | .284 | .314 | .228 | .022 | .020 | .024 | .003 | .026 | .254 |

* p < .05. **p < .01.
Table 13
Paired sample t-tests comparing pairs of initiating and responding variables for Examiner 1 only; Original risk-group criteria

<table>
<thead>
<tr>
<th>Variable Pair</th>
<th>Paired Mean Diff.</th>
<th>t score</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Target Group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IJA-IBR Composites</td>
<td>-.304</td>
<td>-.769</td>
<td>9</td>
<td>.462</td>
</tr>
<tr>
<td>Rate Car IJA-IBR</td>
<td>-.615</td>
<td>-.649</td>
<td>6</td>
<td>.540</td>
</tr>
<tr>
<td>PctCorrRJA/RBR</td>
<td>9.228</td>
<td>.711</td>
<td>7</td>
<td>.500</td>
</tr>
<tr>
<td><strong>Comparison Group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IJA-IBR Composite</td>
<td>.412</td>
<td>1.026</td>
<td>17</td>
<td>.319</td>
</tr>
<tr>
<td>Rate Car IJA-IBR</td>
<td>-.193</td>
<td>-.522</td>
<td>13</td>
<td>.610</td>
</tr>
<tr>
<td>PctCorrRJA-RBR</td>
<td>18.026</td>
<td>2.210</td>
<td>15</td>
<td>.043*</td>
</tr>
</tbody>
</table>

*p < .05 ; **p < .01
Table 14

Paired sample t-tests comparing pairs of initiating and responding variables for Examiner 1 only; Extreme risk-group criteria

<table>
<thead>
<tr>
<th>Variable Pair</th>
<th>Paired Mean Diff.</th>
<th>t score</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High-Risk Group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IJA-IBR Composites</td>
<td>-.589</td>
<td>-1.389</td>
<td>7</td>
<td>.207</td>
</tr>
<tr>
<td>Rate Car IJA-IBR</td>
<td>-.961</td>
<td>-.922</td>
<td>5</td>
<td>.399</td>
</tr>
<tr>
<td>PctCorrRJA/RBR</td>
<td>15.458</td>
<td>.932</td>
<td>5</td>
<td>.394</td>
</tr>
<tr>
<td><strong>Low-Risk Group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IJA-IBR Composite</td>
<td>.631</td>
<td>1.123</td>
<td>9</td>
<td>.290</td>
</tr>
<tr>
<td>Rate Car IJA-IBR</td>
<td>-.227</td>
<td>-.442</td>
<td>7</td>
<td>.672</td>
</tr>
<tr>
<td>PctCorrRJA-RBR</td>
<td>25.925</td>
<td>2.256</td>
<td>8</td>
<td>.054</td>
</tr>
</tbody>
</table>

*p < .05; **p < .01
Figure 1

Room Set-Up for Early Social Communication Scales (ESCS)
Figure 2

Scatterplot of Risk Score and IJA Composite Score delineating the 50th and 90th percentiles for children’s risk scores.
Figure 3

Scatterplot of risk score and rate of total initiating behaviors
Figure 4

Bar graph for interaction effect of Examiner × Risk Group for the variable Rate of Total Frequency of Initiating Behaviors
Figure 5
Risk Score predicting Rate of Initiating Joint Attention, Rate of Total Initiating Behaviors, and Initiating Joint Attention Composite Scores: Examiner 1 Cases
REFERENCES


