8-15-2013

The Utility of the Modified Checklist for Autism in Toddlers in a Preschool-Age Special Education Sample

JoAnna R. Cogan-Ferchalk
Indiana University of Pennsylvania

THE UTILITY OF THE MODIFIED CHECKLIST FOR AUTISM IN TODDLERS IN A PRESCHOOL-AGE SPECIAL EDUCATION SAMPLE

A Dissertation

Submitted to the School of Graduate Studies and Research

in Partial Fulfillment of the

Requirements for the Degree

Doctor of Education

JoAnna R. Cogan-Ferchalk

Indiana University of Pennsylvania

August 2013

Indiana University of Pennsylvania
School of Graduate Studies and Research
Department of Educational and School Psychology

We hereby approve the dissertation of

JoAnna R. Cogan-Ferchalk

Candidate for the degree of Doctor of Education

May 28th, 2013 — Signature on File
Joseph F. Kovaleski, D.Ed.
Professor of Educational and School Psychology, Advisor

May 28th, 2013 — Signature on File
Mark J. Staszkiewicz, Ed.D.
Professor of Educational and School Psychology

May 28th, 2013 — Signature on File
Timothy J. Runge, Ph.D.
Assistant Professor of Educational and School Psychology

May 28th, 2013 — Signature on File
Lisa C. Newell, Ph.D.
Associate Professor of Psychology

ACCEPTED

Signature on File
Timothy P. Mack, Ph.D.
Dean
School of Graduate Studies and Research

Title: The Utility of the Modified Checklist for Autism in Toddlers in a Preschool-Age Special Education Sample

Author: JoAnna R. Cogan-Ferchalk

Dissertation Chair: Dr. Joseph F. Kovaleski

Dissertation Committee Members: Dr. Mark J. Staszkiewicz
Dr. Timothy J. Runge
Dr. Lisa C. Newell

The number of young children with autism is ever-increasing, and school psychologists are more frequently required to identify these students. Valid screening tools are needed in order to focus school psychologists' time on those students in need of intensive evaluations in Early Intervention programs. The purpose of this study was to examine the utility of the Modified Checklist for Autism in Toddlers (M-CHAT; Robins, Fein, & Barton, 1999), an autism screening tool, for preschool-age students who have been identified as needing special education services, using both traditional scoring methods and the newly introduced Best7 scoring procedure. This study also examined which items on the M-CHAT were most associated with students' receiving an educational classification of autism, as discriminated from students with other developmental delays. In addition, this study attempted to determine whether a two-factor solution (i.e., social communication

deficits and unusual behaviors) or a different model best described parent ratings on the M-CHAT.

Overall, findings indicated that the M-CHAT correctly classified 62% of students in the current sample, while incorrectly classifying 38%. Sensitivity was .64 for the current sample while the specificity was .60. The positive predictive value for the M-CHAT was .61 and the negative predictive value was .64. This held true regardless of the scoring method. No differences were found with the use of traditional scoring vs. Best7 scoring methods. It was found that the failure of Question 13 (Does your child imitate you?), and Question 2 (Does your child take an interest in other children?) were the best predictors of students receiving an educational classification of autism.

Results of principal components analyses indicated that the M-CHAT is composed of two components, though the emphasis of these components was different for students with and without autism. While the factors were sufficiently different to prevent combining the groups for a single factor analysis, neither group demonstrated a clear delineation between social communication and unusual behaviors. Thus, the M-CHAT does not appear to be measuring autism as per the newly recommended two-factor model of social communication and unusual behaviors.

ACKNOWLEDGMENTS

This dissertation would not have been completed without support and guidance from many people in my life. I would first like to thank my husband, Matt Ferchalk, who has been my constant source of motivation (and competition!) as we worked together on this never-ending journey for the last 10 years. It hardly seems real that we've reached the end at last. I can't wait to see what our lives beyond dissertation are like.

Second, I would like to thank my committee chairperson, Dr. Joseph Kovaleski, who has patiently guided me through this process and ensured that I reached the other side mostly unscathed. I would also like to thank all of my committee members, whose input and feedback was invaluable to completing this process: Dr. Timothy Runge, Dr. Lisa Newell, and Dr. Mark Staszkiewicz.

My family, as well, deserves thanks for making me who I am today. My stubbornness and refusal to take no for an answer were at times the only reason I continued on with this journey. And of course, I must thank the Ferchalks, the best in-laws a girl could have, whose support and insistence that I would finish someday never wavered.

I must also thank all of my Early Intervention ladies, whose constant inquiries into my progress were a wonderful source of motivation to finish this project. Most especially, I would like to thank Kim Bracey and Sarah Thompson who both spent countless hours tracking down and organizing the necessary archival records for this study. Without their help this research would not have been possible. Further, I would like to thank Sarah Thompson not just for helping with data collection, but for letting me cry, always being willing to listen to me, and showing me the true meaning of friendship when I needed it the most. Your kindness, compassion, and support will never be forgotten.

***

For my angel, gone too soon, and for my little man, who can't come soon enough.

***

TABLE OF CONTENTS

Chapter Page

I INTRODUCTION....................................1

Statement of the Problem.........................5
Significance of the Problem......................9
Research Questions..............................12
Definition of Terms.............................13
Educational Definition.....................13
DSM Definitions..........................14
Overview of Definitions....................25
Assumptions.....................................26
Limitations.....................................27
Delimitations...................................28
Summary...28

II REVIEW OF LITERATURE...........................29

History of Autism...............................29
Early Medical and Mental Health History....................................30
Autism in the *DSM*........................34
Educational History........................36
Analysis of Autism Spectrum Disorder Evaluation Instruments.........................37
Analysis of the Social Responsiveness Scale (SRS)................38
Analysis of the Autism Diagnostic Interview-Revised (ADI-R).................39
Analysis of the Autism Diagnostic Observation Schedule (ADOS)...............40
Analysis of the Autism Spectrum Rating Scales (ASRS).......................41
Summary of Analyses........................42
Autism in Preschool-Age Students................43
Social and Communication Development in Preschool-Age Students.................43
Characteristics of Autism Identifiable in Early Childhood........................44
Autism Identification...........................47
Best Practices in Autism Identification...........................48
Autism Evaluation Instruments..............51
Considerations for Autism Identification in Young Children.........................54

Chapter Page

Necessity for Accurate Autism Screening.....55
Screening Measures for Autism.....56
Level II Screening Measures.....57
Level I Screening Measures.....60
Background and Development of the M-CHAT.....63
Rationale for Study.....71
Summary.....76

III METHOD AND PROCEDURES.....77

Design.....77
Population.....77
Sample.....79
Inclusion Criteria.....79
Exclusion Criteria.....80
Participants.....80
Measurement.....82
Dependent Variable.....82
Independent Variables.....83
Procedure.....87
Data Analysis.....88
Research Question 1.....88
Research Question 2.....90
Research Question 3.....91
Research Question 4.....92
Research Question 5.....93
Summary.....94

IV DATA AND ANALYSES.....98

Research Question 1 Results.....98
Research Question 2 Results.....104
Research Question 3 Results.....109
Research Question 4 Results.....113
Research Question 5 Results.....123
Principal Components Analysis for Students with Autism.....129
Principal Components Analysis for Students without Autism.....135
Conclusions of Principal Components Analyses.....140
Summary.....141

Chapter Page

V DISCUSSION..143

Discussion of Findings.........................144
Research Question 1.......................144
Research Question 2.......................146
Research Question 3.......................148
Research Question 4.......................149
Research Question 5.......................152
Limitations....................................156
Implications for Future Research..............159
Implications for School Psychology............162
Conclusions....................................164
Summary..165

REFERENCES..169

APPENDICES

Appendix A, M-CHAT Protocol........................188

Appendix B, M-CHAT Scoring Instructions............189

Appendix C, Permission to Duplicate the M-CHAT.....190

Appendix D, Permission to Access Archival Data.....191

LIST OF TABLES

Table Page

1 Sex, Age, and Race of Participants.................81

2 Comparison Group Educational Classifications.......82

3 Source of Autism Diagnoses.........................87

4 Research Questions, Hypotheses, Variables, Statistical Analyses, and Assumptions...............95

5 Pass/Fail Rates on the M-CHAT Using Traditional Scoring...............................100

6 ROC Results for Critical Items and All Items on the M-CHAT.....................................102

7 Pass/Fail Rates on the M-CHAT Using Best7 Scoring...106

8 ROC Results for Best7 Items and All Items on the M-CHAT.....................................107

9 Comparison of Pass/Fail Rates on the M-CHAT for Traditional vs. Best7 Scoring Procedures......................................111

10 Comparison of Sensitivity, Specificity, Positive Predictive Values, and Negative Predictive Values for Traditional vs. Best7 Scoring Procedures...............................111

11 Comparison of the AUCs for Traditional vs. Best7 Scoring Procedures.........................112

12 VIF and Tolerance Statistics for M-CHAT Questions.......................................115

13 Point Biserial Correlation Matrix for M-CHAT Questions.......................................116

14 Wilks's Lambda for Stepwise Discriminant Function Analysis................................119

Table Page

15 M-CHAT Performance with Varying Cutoff Criteria...121

16 Point Biserial Correlation Matrix for M-CHAT Questions for Students with Autism.........125

17 Point Biserial Correlation Matrix for M-CHAT Questions for Students without Autism..127

18 Actual and Randomly Generated Eigenvalues Used for Parallel Analysis for Students with Autism..131

19 Total Variance Explained for Principal Components Analysis Using a Two-Factor Solution for Students with Autism.................133

20 Rotated Component Matrix for Principal Components Analysis for Students with Autism......134

21 Actual and Randomly Generated Eigenvalues Used for Parallel Analysis for Students without Autism....................................137

22 Total Variance Explained for Principal Components Analysis Using a Two-Factor Solution for Students without Autism..............138

23 Rotated Component Matrix for Principal Components Analysis for Students without Autism..139

LIST OF FIGURES

Figure Page

1 ROC curve including the six critical items on the M-CHAT....................................103

2 ROC curve including all 23 items on the M-CHAT..103

3 ROC curve including the Best7 items on the M-CHAT..108

4 ROC curve including all 23 items on the M-CHAT..108

5 Box plot depicting pass/fail rates for question 16 on the M-CHAT for students with and without ASD (autism spectrum disorder), indicating that only three students failed this question..........118

6 Scree plot of eigenvalues for students with Autism indicating significant change of slope after two factors.................................130

7 Scree plot of eigenvalues for students without Autism indicating significant change of slope after two factors.................................136

CHAPTER I

INTRODUCTION

Among the most critical issues facing educators today is the rapid increase of the number of children with autism. Just a few years ago autism was reported to occur at a rate of 1 in 150 children (Layne, 2007); however in its most recent update, the Centers for Disease Control and Prevention estimated that autism now occurs in 1 of every 50 school-aged children in the United States (Blumberg et al., 2013). With this increase school evaluation teams are called upon to accurately identify autism and recommend appropriate services for these children (Wilkinson, 2011). This is especially important for Early Intervention (EI) programs, which provide special education services for preschool-age children (ages 3-5) in many states, as this is the age range when most cases of autism are first identified (Shattuck et al., 2009). Accurate means of screening for and identifying autism are necessary to ensure that these children receive appropriate services. Early identification is vital, as research has shown that when interventions are provided at an early age deficits related to autism can be altered (Crane & Winsler, 2008; Levy et al., 2007; National Research Council, 2001; Robins & Dumont-Mathieu, 2006).

The following scenario emphasizes the importance of accurately identifying students with autism: A school psychologist was asked to come into a preschool special education classroom to consult with a teacher regarding a student with significant behavior problems. The student was previously evaluated and found eligible for special education services under the educational classification of developmental delay. The school psychologist's observation revealed that the student became agitated during group activities and attempted to leave the situation. The student's language was limited and he lacked a means to request his wants and needs or to protest things he did not want, which led to disruptive behaviors caused by frustration. The student actively avoided interactions with the other children in the classroom, but continuously sought out teacher attention. An interview with the teacher revealed similar concerns. The teacher was unable to motivate the student to participate in group activities involving other students and she needed to have a staff person with him constantly.

A functional behavior assessment (FBA) revealed a very specific pattern. The problem behaviors identified were related to the student's limited language, delayed social skills, difficulty with transitions between activities, and an inability to tolerate the sensory input associated with participating in group activities. In reviewing this student's case it became clear that the student was not simply a child with a developmental delay and accompanying behavior, but rather his difficulties were most likely related to an autism spectrum disorder (ASD). After compiling this information, a reevaluation was undertaken and the student was identified under the educational classification of autism. After appropriate accommodations and supports were introduced into the classroom and consultation with the teacher provided her with an understanding of the student's strengths and needs, the problem behaviors significantly decreased and participation increased.

In the above situation, the student was screened for autism using the Modified Checklist for Autism in Toddlers (M-CHAT; Robins, Fein, & Barton, 1999), which is a commonly used screener for autism. The student passed the screener; therefore a school psychologist was not involved in his initial evaluation process and concerns relating to autism were not addressed until problem behaviors were reported, nearly a year later.

This is not an isolated situation. Although Pennsylvania's Chapter 14 Special Education Regulations mandate that school psychologists are included when students are being evaluated for autism (2008, §14.123), their inclusion is not mandated by these regulations for evaluations of students suspected of having a developmental delay, a category specific to EI programs (§14.153). Developmental delay is a broad category, which includes demonstrating a 25% delay on a developmental assessment when using chronological age equivalents, or earning a standard score that falls "1.5 standard deviations below the mean" (§14.101).

The vast majority of children with autism would meet these criteria; therefore there is little risk of a student not qualifying for needed services. Yet, as demonstrated above, it is vital for students with autism to be accurately identified as such, in order to provide appropriate services and accommodations as part of their Individualized Education Plan (IEP). In Pennsylvania it is common for a school psychologist to be part of an evaluation if autism is suspected; however, if autism or other significant disabilities are not suspected, the entire evaluation can be completed without input from a school psychologist. This begs the question: Without adequate screening measures in place, how can it be determined when a more significant disability, such as autism, may be present and require more intensive evaluation measures? Students may pass autism screeners such as the M-CHAT, even when concerns are present. While students will qualify for and receive special education services, if their true disability has not been accurately identified, their IEPs will not address all areas of need and their teachers may not have the necessary information and tools to adequately support them.

Statement of the Problem

There is a clear need for accurate screening measures to identify potential cases of autism in preschool-age special education programs. Students who are referred for evaluation though EI programs are routinely screened for autism using the M-CHAT and similar measures, yet it is possible for students to pass these screeners, leaving evaluation teams unaware of larger issues. The M-CHAT is one of the most extensively used screeners for autism because it is free and readily available to schools, pediatricians, and mental health agencies (Robins, 2008). This leads to an important research question. Is the M-CHAT accurately identifying students in need of further evaluation for autism?

The purpose of this study is to examine the utility of the M-CHAT for preschool-age special education students.

This study also seeks to determine which items on the M-CHAT are most associated with students' receiving an educational classification of autism, as discriminated from students with other developmental delays. In addition, this study seeks to determine whether a two-factor model of autism identification (i.e., social communication deficits and unusual/repetitive behaviors) or a different model best describes parent ratings on the M-CHAT.

The M-CHAT is a 23-item, parent-completed checklist designed to identify young children in need of further evaluation for autism (Robins et al., 1999; see Appendix A). Items on the checklist are classified as critical or noncritical items. Based on discriminant function analysis, Robins, Fein, Barton, and Green (2001) recommended that children who failed two or more critical items or any three items should be referred for further evaluation (see Appendix B). In 2010 Robins et al. presented preliminary findings regarding a new scoring algorithm for the M-CHAT. These scoring criteria, known as the Best7, identified seven critical items rather than the previous six (see Appendix B).

Currently, the vast majority of research focuses on the use of the M-CHAT for autism screening for children aged 16-30 months (Dumont-Mathieu & Fein, 2005; Kleinman et al., 2008; Mawle & Griffiths, 2006; Pandey et al., 2008; Robins, 2008), the age range it was intended to screen and the age range examined in the original validation study (Robins et al., 2001). Overall, it has been found to be a useful tool for referring young children in need of autism evaluations (Robins, 2008; Robins et al., 2001).

The M-CHAT is also widely used for students older than 30 months, as it a free and easily accessible screener (Robins, 2008); however, there has been relatively little research examining the M-CHAT for students above the intended age range (Eaves, Wingert, & Ho, 2006; Snow & Lecavalier, 2008) despite its frequent use in this population. Of the research available, little information pertained to whether particular items or patterns of responding were more highly associated with children with autism (Eaves et al., 2006; Robins, et al., 2008). Further, of the studies that examined the M-CHAT's use for older students, all took place in clinical or university settings. An electronic journal article search yielded no studies in the published, peer-reviewed literature examining the M-CHAT's use as a screening tool in educational settings. This study seeks to add to the literature that examines the use of the M-CHAT for preschool age students in special education programming.

Historically, instruments used to identify children with autism have focused on three main areas of concern: Communication, social interaction, and restricted/repetitive behaviors; based on the recommendations of the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders*' (*DSM-IV-TR*; American Psychiatric Association [APA], 2000). However, the newer conceptualizations of autism do not support the three-factor model (APA, 2012, 2013). Through the process of creating and validating the Autism Spectrum Rating Scales (ASRS; Goldstein & Naglieri, 2010), Goldstein and Naglieri (2011) found that a two-factor model best supported the identification of autism in children under the age of 6. Based on factor analysis, the authors proposed a combined social communication factor along with an unusual behaviors factor. Similarly, studies examining the use of the Autism Diagnostic Observation Schedule (ADOS; Lord, Rutter, DiLavore, & Risi, 1999) have found that the use of a two-factor scoring algorithm including social affect and restricted, repetitive behaviors resulted in better predictive validity (Gotham et al., 2008; Gotham, Risi, Pickles, & Lord, 2007; Oosterling et al., 2010).

The APA also recently revised its recommendations for the identification of autism spectrum disorders. Based on the ongoing research conducted by its members and focus workgroups, the APA has made the recommendation for a two-factor model as well, identified as social communication deficits and "restricted, repetitive patterns of behaviors" (APA, 2013, p. 50). Analysis of student performance on the M-CHAT would also in part seek to determine whether these factors or a different pattern of factors are measured by the M-CHAT.

Significance of the Problem

The increasing prevalence of autism represents a growing issue in autism assessment. As prevalence rates are steadily increasing, the need for effective autism identification in the schools becomes more and more important.

Research indicates that in the majority of cases, identification of autism is the responsibility of school systems. Glascoe (2000) found that of children with autism diagnoses, 70% had been identified by schools and educational systems, as opposed to 30% of children who received their diagnoses from mental health agencies, medical practitioners, or other sources. Comparable findings were reported by Palfrey, Singer, Walker, and Butler (1987), who determined that nearly 80% of children had been identified by educational personnel, as opposed to other agencies. Further, nearly 75% of children with autism were reported to have been identified first through the school system in a 2003 study (Yeargin-Allsopp et al.), as opposed to receiving mental health or medical diagnoses. These findings indicate that much of the burden of screening and identifying autism is falling on schools, as on average only 20-30% of children are diagnosed with autism by other medical or mental health providers prior to their receiving an educational evaluation. An adequate screening process is vital to ensuring that these children receive the appropriate educational classification and services.

Additionally, as autism prevalence rates increase, so does the prevalence of due process hearings relating to autism (Ikeda, 2002; Noland & Gabriels, 2004). Federal special education law states that children are entitled to "full, individualized and appropriate educational evaluations by the Local Education Agency (LEA) designed to identify all of the child's special education needs" (Individuals with Disabilities Education Improvement Act [IDEIA], as cited in Noland & Gabriels, 2004, p. 267). This includes making appropriate decisions regarding a child's educational classification and placement.

In a review of 45 due process cases surrounding autism Yell and Drasgow (2000) found five main reasons why parents initiated legal proceedings with school districts (see also Noland & Gabriels, 2004). Two of these reasons directly relate to evaluation issues: The first being that the school district did not evaluate all areas of need, and the second being that the school district conducted evaluations that did not actually assess autism. Additional reasons for due process hearings were that the school district did not have personnel trained in autism, inadequate IEPs were developed, and parents were not involved in the IEP process (Yell & Drasgow, 2000). These issues indirectly relate to evaluation, as IEPs naturally address needs identified in the evaluation process. In addition, due process cases tend to be extremely time consuming and expensive for school districts, students, and parents. Thus, proper identification and evaluation of students with autism is needed to avoid these confrontations.

Overall, the number of children with autism in the United States continues to increase steadily. Schools, and more specifically EI programs, are being called upon to accurately identify these students and provide appropriate special education services. Accurate screening for autism is a necessity in order to ensure that these evaluations take place.

Research Questions

1. How does the M-CHAT perform as an autism screening measure in the current sample of preschool-age students (30-72 months) based on the traditional scoring procedures established for 16-30 month olds?
2. How does the M-CHAT perform as an autism screening measure in the current sample of preschool-age students (30-72 months) based on the Best7 scoring procedures suggested for 16-30 month olds?
3. What are the differences in the M-CHAT's performance as an autism screening measure in the current sample, based on the use of the traditional scoring procedures vs. the Best7 scoring procedures?
4. Which items on the M-CHAT are most predictive of preschool-age students (30-72 months) in the current sample receiving an educational classification of autism?
5. Does a principal components analysis of parent ratings indicate that the M-CHAT is measuring two components of autism (i.e., social communication deficits and unusual/repetitive behaviors), as per the newly recommended autism identification model, or will a

different solution best represent findings from the current sample?

Definition of Terms

Autism spectrum disorders are defined differently by educational law and the mental health field. It is important to note these distinctions, though for the purposes of this study, both the educational and mental health definitions of autism spectrum disorders will be referred to with the inclusive label of autism.

Educational Definition

In the IDEIA Federal Regulations (2004) autism is defined as follows:

> **Autism.**
>
> (i) Autism means a developmental disability significantly affecting verbal and nonverbal communication and social interaction, generally evident before age three, that adversely affects a child's educational performance. Other characteristics often associated with autism are engagement in repetitive activities and stereotyped movements, resistance to environmental change or change in daily routines, and unusual responses to sensory experiences.

(ii) Autism does not apply if a child's educational performance is adversely affected primarily because the child has an emotional disturbance, as defined in paragraph (c)(4) of this section.

(iii) A child who manifests the characteristics of autism after age three could be identified as having autism if the criteria in paragraph (c)(1)(i) of this section are satisfied (§300.8).

DSM **Definitions**

In the *DSM-IV-TR*, autism spectrum disorders fall under the umbrella term of pervasive developmental disorders (PDD; APA, 2000). These include autistic disorder; Asperger's disorder; pervasive developmental disorder, not otherwise specified (PDD-NOS); Rett's disorder; and childhood disintegrative disorder. Rett's disorder will be addressed in a later section (see Overview of Definitions). The remaining diagnoses are defined as follows:

Autistic disorder.

A. A total of six (or more) items from (1), (2), and (3), with at least two from (1), and one each from (2) and (3):

(1) qualitative impairment in social interaction, as manifested by at least two of the following:

(a) marked impairment in the use of multiple nonverbal behaviors such as eye-to-eye gaze, facial expression, body postures, and gestures to regulate social interaction

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

(c) a lack of spontaneous seeking to share enjoyment, interests, or achievements with other people (e.g., by a lack of showing, bringing, or pointing out objects of interest)

(d) lack of social or emotional reciprocity

(2) qualitative impairments in communication as manifested by at least one of the following:

(a) delay in, or total lack of, the development of spoken language (not accompanied by an attempt to compensate through alternative modes of communication such as gesture or mime)

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

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

(d) lack of varied, spontaneous make-believe play or social imitative play appropriate to developmental level

(3) restricted repetitive and stereotyped patterns of behavior, interests, and activities, as manifested by at least one of the following:

(a) encompassing preoccupation with one or more stereotyped and restricted patterns of interest that is abnormal either in intensity or focus

(b) apparently inflexible adherence to specific, nonfunctional routines or rituals

(c) stereotyped and repetitive motor mannerisms (e.g., hand or finger flapping or twisting, or complex whole body movements)

(d) persistent preoccupation with parts of objects

B. Delays or abnormal functioning in at least one of the following areas, with onset prior to age 3 years: (1) social interaction, (2) language as used in social communication, or (3) symbolic or imaginative play.

C. The disturbance is not better accounted for by Rett’s Disorder or Childhood Disintegrative Disorder (APA, 2000, p. 75).

Asperger's disorder.

A. Qualitative impairment in social interaction, as manifested by at least two of the following:

(1) marked impairment in the use of multiple nonverbal behaviors such as eye-to-eye gaze, facial expression, body postures, and gestures to regulate social interaction

(2) failure to develop peer relationships appropriate to developmental level

(3) a lack of spontaneous seeking to share enjoyment, interests, or achievements with other people (e.g., by a lack of showing, bringing, or pointing out objects of interest to other people)

(4) lack of social or emotional reciprocity

B. Restricted repetitive and stereotyped patterns of behavior, interests, and activities, as manifested by at least one of the following:

(1) encompassing preoccupation with one or more stereotyped and restricted patterns of interest that is abnormal either in intensity or focus

(2) apparently inflexible adherence to specific, nonfunctional routines or rituals

(3) stereotyped and repetitive motor mannerisms

(e.g., hand or finger flapping or twisting, or complex whole-body movements)

(4) persistent preoccupation with parts of objects

C. The disturbance causes clinically significant impairment in social, occupational, or other important areas of functioning.

D. There is no clinically significant general delay in language (e.g., single words used by age 2 years, communicative phrases used by age 3 years).

E. There is no clinically significant delay in cognitive development or in the development of age-appropriate self-help skills, adaptive behavior (other than in social interaction), and curiosity about the environment in childhood.

F. Criteria are not met for another specific Pervasive Developmental Disorder or Schizophrenia (APA, 2000, p. 84).

Pervasive developmental disorder not otherwise specified (PDD-NOS).

This category should be used when there is a severe and pervasive impairment in the development of reciprocal social interaction or verbal and nonverbal communication skills, or when stereotyped behavior, interests, and activities are present, but the criteria are not met for a specific pervasive developmental disorder, schizophrenia, schizotypal personality disorder, or avoidant personality disorder. For example, this category includes "atypical autism" -presentations that do not meet the criteria for autistic disorder because of late age of onset, atypical symptomatology, or subthreshold symptomatology, or all of these (APA, 2000, p. 84).

Childhood disintegrative disorder.

A. Apparently normal development for at least the first 2 years after birth as manifested by the presence of age-appropriate verbal and nonverbal communication, social relationships, play, and adaptive behavior.

B. Clinically significant loss of previously acquired skills (before age 10 years) in at least two of the following areas:

(1) expressive or receptive language

(2) social skills or adaptive behavior

(3) bowel or bladder control

(4) play

(5) motor skills

C. Abnormalities of functioning in at least two of the following areas:

(1) qualitative impairment in social interaction (e.g., impairment in nonverbal behaviors, failure to develop peer relationships, lack of social or emotional reciprocity)

(2) qualitative impairments in communication (e.g., delay or lack of spoken language, inability to initiate or sustain a conversation, stereotyped and repetitive use of language, lack of varied make-believe play)

(3) restricted, repetitive, and stereotyped patterns of behavior, interests, and activities, including motor stereotypies and mannerisms

D. The disturbance is not better accounted for by another specific Pervasive Developmental Disorder or by Schizophrenia (APA, 2000, p. 79).

Autism spectrum disorder. The above definitions were the current recommendations in the mental health field for differentiating autism spectrum disorders at the time data were collected for the present study. As mentioned previously, however, criteria have been recently revised for the new edition of the *DSM*, to allow for a two-factor model of autism. In addition, it is currently recommended that all autism diagnoses fall under the inclusive term autism spectrum disorder. Per the APA's *DSM-5* development website:

> Differentiation of autism spectrum disorder from typical development and other "nonspectrum" disorders is done reliably and with validity; while distinctions among disorders have been found to be inconsistent over time, variable across sites and often associated with severity, language level or intelligence rather than features of the disorder.
>
> Because autism is defined by a common set of behaviors, it is best represented as a single diagnostic category that is adapted to the individual's clinical presentation by inclusion of clinical specifiers (e.g., severity, verbal abilities and others) and associated features (e.g., known genetic disorders, epilepsy, intellectual disability and others). A single spectrum disorder is a better reflection of the state of knowledge about pathology and clinical presentation; previously, the criteria were equivalent to trying to "cleave meatloaf at the joints" (APA, 2012).

The following is the newest definition of autism spectrum disorder, included in the *DSM-5* when it was published in May 2013:

A. Persistent deficits in social communication and social interaction across multiple contexts, as manifested by the following, currently or by history (examples are illustrative, not exhaustive; see text):

   1. Deficits in social-emotional reciprocity, ranging, for example, from abnormal social approach and failure of normal back-and-forth conversation; to reduced sharing of interests, emotions, and affect; to failure to initiate or respond to social interactions.
   2. Deficits in nonverbal communicative behaviors used for social interaction, ranging, for example, from poorly integrated verbal and nonverbal communication; to abnormalities in eye contact and body language or deficits in understanding and use of gestures; to a total lack of facial expressions and nonverbal communication.
   3. Deficits in developing, maintaining, and understanding relationships, ranging, for

example, from difficulties adjusting behavior to suit different social contexts; to difficulties in sharing imaginative play or in making friends; to absence of interest in peers.

Specify current severity:

Severity is based on social communication impairments and restricted, repetitive patterns of behavior...

B. Restricted, repetitive patterns of behavior, interests, or activities, as manifested by at least two of the following, currently or by history (examples are illustrative, not exhaustive; see text):

1. Stereotyped or repetitive motor movements, use of objects, or speech (e.g., simple motor stereotypies, lining up toys or flipping objects, echolalia, idiosyncratic phrases).
2. Insistence of sameness, inflexible adherence to routines, or ritualized patterns of verbal or nonverbal behavior (e.g., extreme distress at small changes, difficulties with transitions, rigid thinking patterns, greeting rituals, need to take same route or eat same food every day).

3. Highly restricted, fixated interests that are abnormal in intensity or focus (e.g., strong attachment to or preoccupation with unusual objects, excessively circumscribed or perseverative interests).

4. Hyper- or hyporeactivity to sensory input or unusual interest in sensory aspects of environment (e.g., apparent indifference to pain/temperature, adverse response to specific sounds or textures, excessive smelling or touching of objects, visual fascination with lights or movement).

Specify current severity:

Severity is based on social communication impairments and restricted, repetitive patterns of behavior...

C. Symptoms must be present in the early developmental period (but may not become fully manifest until social demands exceed limited capacities, or may be masked by learned strategies in later life).

D. Symptoms cause clinically significant impairment in social, occupational, or other important areas of current functioning.

E. These disturbances are not better explained by intellectual disability (intellectual developmental disorder) or global developmental delay. Intellectual disability and autism spectrum disorder frequently co-occur; to make comorbid diagnoses of autism spectrum disorder and intellectual disability, social communication should be below that expected for general developmental level (APA, 2013, pp. 50-51).

Overview of Definitions

Previously, educational and mental health definitions of autism were somewhat varied; however, with the new definition of autism spectrum disorder they are more closely aligned, each with a generally inclusive term to describe all spectrum disorders. In the new mental health model the diagnoses of autistic disorder, Asperger's disorder, PDD-NOS, and childhood disintegrative disorder are subsumed by the diagnosis of autism spectrum disorder, as research indicates that these terms cannot be reliably differentiated from each other (APA, 2012, 2013). The diagnosis of Rett's disorder is not included in the *DSM-5*, as it is a genetic disorder and is now considered a medical diagnosis rather than a mental health disorder.

For the purposes of this study, autism refers to the federal special education classification of autism, which describes students who display delays in communication and social skills, along with repetitive behaviors and possible difficulties with transitions and changes in routine (IDEIA, 2004). Autism spectrum disorders including autistic disorder, Asperger's disorder, PDD-NOS, and childhood disintegrative disorder will also be included under the general label of autism for the purposes of this paper and will not be referred to separately. The terms autism and autism spectrum disorders will be used synonymously. Rett's disorder will not be included in the term autism in this study as it was removed from the *DSM-5* (APA, 2013). No distinctions will be made between high and low functioning students with autism, as regardless of ability or language levels all students who meet federal special education criteria receive the label of autism.

Assumptions

The following are assumed to be true for the purposes of this study:

1. Parents had sufficient knowledge to read, understand, and correctly interpret questions on the M-CHAT.

2. Parents who rated their children using the M-CHAT had sufficient knowledge of their child to provide accurate ratings.
3. Parents provided valid ratings that honestly reflected their child's functioning.
4. Records accessed contained accurate data regarding participants' scores on the M-CHAT, demographic information, and educational classification.
5. Participants were identified under the correct educational classification (i.e., autism or other developmental delays).

Limitations

A major assumption of this study is that students have been identified under the correct educational classification. It is possible, however, that some students with autism may not have been correctly identified, especially given that the current study is questioning whether the M-CHAT is an appropriate autism screening measure for preschool-age students. If the M-CHAT is not accurately recognizing students in need of further testing for autism there may be a small number of students who are classified incorrectly as having a developmental delay only.

Delimitations

This study will examine a sample from one intermediate unit located in Eastern/Central Pennsylvania. This may limit the population to which the findings can be generalized.

Summary

The number of children with autism is ever-increasing, and school psychologists are more frequently required to identify these students. Valid screening tools are needed in order to focus school psychologists' time on those students in need of intensive evaluations in EI programs. This study seeks to examine preschool-age students' performance on the M-CHAT, based on parent ratings, in order to determine how many students pass and fail the M-CHAT and whether these students receive an educational classification of autism. In addition, this study seeks to determine whether a two-factor model of autism identification (i.e., social communication deficits and unusual/repetitive behaviors) or a different model best describes parent ratings on the M-CHAT.

CHAPTER II

REVIEW OF LITERATURE

This chapter will provide an overview of the history of autism and the current changes in the understanding and definitions of autism. This chapter will also provide a review of characteristics of autism that are observable in preschool-age students, and the recommended model of autism identification in Pennsylvania. In addition, the development of autism screening tools, including the Modified Checklist for Autism in Toddlers (M-CHAT; Robins, Fein, & Barton, 1999), will be explored, along with a review of pertinent research and the rationale for the proposed research study.

History of Autism

Descriptions of children and adults with characteristics similar to our modern understanding of autism have been noted for centuries (Frith, 2003; Gillberg, 1998; Hippler & Klicpera, 2003; Wing, 1997). Early on, some of these children were thought to be psychotic, insane, suffering from childhood forms of schizophrenia, mentally retarded (currently termed intellectually disabled), or even demon possessed (APA, 1952, 1968; Wing, 1997, 2005). Despite the case studies and anecdotes of individuals who all bore similar characteristics, the idea that these cases constituted a unique diagnostic category, eventually to be called autism, was not suggested in the literature until the mid-20th century.

Early Medical and Mental Health History

It was not until 1943 that Leo Kanner, director of child psychiatry at the Johns Hopkins Hospital in Maryland, first described 11 case studies of children all demonstrating similar characteristics that he termed "autistic disturbances of affective contact" (p. 217). These children demonstrated features including difficulty coping with change, echolalic language that did not serve the purpose of social communication, and perseveration on organization or knowledge about details that would not interest most others. Kanner believed that "these characteristics form a unique 'syndrome,' not heretofore reported, which seems to be rare enough, yet is probably more frequent than is indicated by the paucity of observed cases" (p. 242). This was the first instance of the term autism being used in published literature to describe these features in children (Wing, 1997).

Kanner (1943) noted that many of these children had been labeled as "feebleminded" or "childhood schizophrenics," which he felt did not adequately describe their unique characteristics (p. 242). Many of the children in Kanner's case studies in fact tested within normal limits on intelligence measures, but struggled with social interaction and conversation. He noted that the children in his case studies represented a wide range of functioning. Some acquired language at the appropriate time or after a delay, but demonstrated use of language that was often restricted to labeling or naming objects, or repeating seemingly inane information such as a page from an encyclopedia from rote memory. Others did not acquire language at all.

Contemporary to Kanner (1943), Hans Asperger, an Austrian psychiatrist, also observed children with similar characteristics. In 1944 Asperger published a collection of case studies of children who demonstrated what he termed "autistic psychopathy" (Asperger, as cited in Hippler & Klicpera, 2003, p. 291). "Asperger believed that [autistic psychopathy] was a continuum disorder merging into the 'normal' continuum, that is, a group of eccentric, withdrawn, but often highly gifted, individuals who manage social integration despite their somewhat odd social interaction or communication" (Hippler & Klicpera, p. 291). Asperger theorized that his autistic psychopathy was similar to Kanner's descriptions of autism, but that the two conditions constituted separate diagnoses and epidemiology (Hippler & Klicpera). His accounts, unfortunately, were not translated and disseminated into English language publications until several decades later (Wing, 1997).

In the 1960s, many researchers attempted to describe and categorize features resembling autism (Wing, 2005). Although the term autism had been introduced some 20 years earlier, these characteristics were originally labeled as childhood schizophrenia (APA, 1968; Wing, 1997, 2005). Researchers eventually disentangled the term autism from schizophrenia when it was observed that children who demonstrated an early onset of symptoms (before 3 years of age) appeared to have a different disorder and outcomes than those who had an onset of 5 years of age or later (Kolvin, 1971; Kolvin, Ounsted, Humphrey, & McNay, 1971).

In the 1970s, an epidemiological study conducted by Wing and Gould (as cited in Wing, 2005) suggested that many individuals demonstrated three core characteristics of autism known as "the triad of impairments" (p. 586; defined as social impairments, delayed language, and restricted/repetitive behaviors) and also introduced the idea that autism occurred in a continuum. This continuum included both Kanner's and Asperger's descriptions of autism, with each of their observations covering just one small part of this continuum. The term continuum was eventually changed to spectrum, which "included the most severe to the subtlest manifestations of the triad" (Wing, 2005, p. 586).

In 1981, Wing published a seminal article that introduced the term Asperger's disorder into the general nomenclature, in reference to individuals who demonstrated high functioning forms of autism. Although the idea that autism was in fact a spectrum disorder had been suggested in 1979 (Wing, 2005), different diagnostic terms continued to be suggested to describe the various levels of functioning within this spectrum, including most notably, Asperger's disorder (Wing, 1981). Thus began decades of debates over whether there were multiple discrete disorders, each with similar characteristics, or whether there was simply one heterogeneous group of autism spectrum disorders (Wilkinson, 2010; Wing, 1981, 1997). The proposed diagnoses and subtypes of autism underwent multiple revisions during various editions of the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM*).

Autism in the *DSM*

In the first and second editions of the *DSM*, autism was not included as its own diagnostic category. Instead, references to autistic-like or withdrawn behaviors were included under the definition of childhood schizophrenia (APA, 1952, 1968). Pervasive developmental disorder (PDD) was first included as a diagnostic category in the third edition of the *DSM* (APA, 1980). The *DSM-III* introduced the term PDD as an umbrella category, which encompassed several subtypes. These included infantile autism, infantile autism residual state, childhood onset pervasive developmental disorder, childhood onset pervasive developmental disorder residual state, and atypical pervasive developmental disorder. These subtypes were meant to help clarify the age of onset as well as symptom severity; however, they led to much diagnostic confusion and few children received accurate autism diagnoses (Factor, Freeman, & Kardash, 1989; Waterhouse, Wing, Spitzer, & Siegel, 1992).

When the *DSM-III* was revised (APA, 1987), the umbrella category of PDD was simplified to include only two diagnoses: Autistic disorder and pervasive developmental disorder not otherwise specified (PDD-NOS). The *DSM-III-R* (1987) also included more concrete and observable criteria to aid with diagnostic decision-making and the criteria for autistic disorder were relaxed, leading to an increased number of diagnoses (Factor et al., 1989; Waterhouse et al., 1992).

The fourth edition of the *DSM* (1994) again sought to expand the diagnoses described under the umbrella term of PDD. The *DSM-IV* incorporated five separate diagnoses including autistic disorder, Asperger’s disorder, Rett's disorder, childhood disintegrative disorder, and PDD-NOS. These diagnostic categories did not change between the *DSM-IV* and the updated *DSM-IV-TR* (APA, 1994, 2000). Concerns arose with this overly complex system of diagnoses. In particular, it was suggested that distinguishing between the various diagnoses led to diagnostic confusion, and the separate diagnoses did not significantly contribute to intervention planning (Wilkinson, 2010; Wing, 2005).

Finally, with the most recent edition of the *DSM*, the APA (2013) proposed to settle the debate and has recommended one overarching diagnosis of autism spectrum disorder in the newly released *DSM-5*. The APA's autism task force concluded that distinguishing between the various diagnoses cannot be done reliably and validly (APA, 2012). All forms of autism, including Asperger's disorder, autistic disorder, and PDD-NOS were subsumed by the term autism spectrum disorder when the *DSM-5* was published in May of 2013.

Educational History

The educational history of autism is somewhat less complex. The Education of All Handicapped Children Act (EHA; PL 94-142), introduced in 1975, mandated a free and appropriate education for all children with disabilities. Although autism was not mentioned specifically in the original EHA or in its 1986 reauthorization (PL 99-457), students with autism could receive special education services provided they met criteria for another disability category, such as mental retardation (now intellectual disability) or speech or language impairment.

Ten years after autism first appeared in the *DSM-III* (APA, 1980), the Individuals with Disabilities in Education Act (IDEA; PL 101-476, 1990), a revision of the EHA, was passed by the United States congress. The IDEA included autism as a federally recognized disability category in education for the first time (Noland & Gabriels, 2004). Autism has remained an educational disability category since 1990, and the definition has not changed substantially in either the 1997 or 2004 reauthorizations of IDEA.

Autism continues to be defined educationally as a "developmental disability significantly affecting verbal and nonverbal communication and social interaction ... that adversely affects a child's educational performance" (§300.8). Repetitive or stereotyped behaviors and unusual responses to sensory experiences may also be present. The educational definition of autism remains sufficiently loose so that any autism spectrum disorder, as defined by the *DSM-IV-TR* (APA, 2000), could be included should the child demonstrate the need for specially designed instruction (Shriver, Allen, & Matthews, 1999). When the *DSM-5* (APA, 2013) was published and the term autism spectrum disorder was formally adopted, the educational definition of autism continues to be sufficiently broad to include any child who meets criteria for an autism spectrum disorder, provided they demonstrate the need for specially designed instruction.

Analysis of Autism Spectrum Disorder Evaluation Instruments

The earliest descriptions of autism stated that "social aloofness and resistance to change in repetitive routines that were *elaborate* in form were the essential diagnostic criteria. If these two were present ... the rest would also be found" (Kanner & Eisenberg, as cited in Wing, 2005, p. 584). Others disagreed with this twopronged approach to autism identification, and suggested that language impairments were an essential part of the diagnosis of autism (Rutter, 1978). Rutter advocated for three criteria when identifying autism: Difficulties with social relationships, deficiencies of language and pre-language, and patterns of stereotyped behaviors and routines.

Based on these arguments, along with descriptions in the *DSM-IV* and *DSM-IV-TR* (APA 1994, 2000), autism has traditionally been viewed as consisting of three distinct factors: Impairment in social interaction, impairment in communication, and the presence of repetitive and stereotypic patterns of behavior. However, the body of research built over the last decade does not support this traditional three-factor model, and instead is leaning toward a two-factor model of autism, as Kanner and Eisenberg (as cited in Wing, 2005) originally suggested. This research consists of factor analyses, cluster analyses, and principal components analyses conducted with a variety of diagnostic autism evaluation instruments, as described below.

Analysis of the Social Responsiveness Scale (SRS)

Constantino et al. (2004) conducted a factor analysis of parent and teacher ratings from the SRS, based on a sample of 226 children aged 4-18 years (Constantino & Gruber, 2005). Parent and teacher ratings were found to be highly correlated with each other, and there were no differences between the factor loadings. The authors concluded that a single unnamed factor explained 35% of the variance, which included all three domains typically identified with autism (language impairments, social deficits, and repetitive/stereotyped behaviors). In addition, multiple unnamed smaller factors were identified, each representing less than 5% of the variance for the rest of the analysis. Separate factors for social and communication domains were not supported.

Analysis of the Autism Diagnostic Interview-Revised (ADI-R)

Constantino et al. (2004) also conducted cluster analysis of the ADI-R (Rutter, Le Couteur, & Lord, 2003) and found that two factors explained 27% of the variance, with 8 additional factors each contributing less than 6% to the final solution. Significant overlap between factors was noted. The first cluster included social deficits, non-verbal communication, and verbal communication. The second cluster included difficulties associated with the three traditional areas of autism deficits (communication, social interaction, restricted/repetitive behaviors). A follow-up principal components analysis revealed 14 factors, with one primary factor explaining 40% of the variance and the remaining factors each contributing a small portion. The primary factor identified again included all three domains traditionally associated with autism (language impairments, social deficits, and repetitive/stereotyped behaviors), which were not differentiated from each other in any way. Overall the authors concluded that there was no evidence for independent subdomains related to autism.

Analysis of the Autism Diagnostic Observation Schedule (ADOS)

Robertson, Tanguay, L'Ecuyer, Sims, and Waltrip (1999) conducted a factor analysis of children's performance on the ADOS and found that a three-factor solution accounted for 70% of the variance. The factors identified were joint attention, affective reciprocity, and theory of mind (ability to engage in symbolic play). Again, this study did not support separate social and language factors.

Gotham, Risi, Pickles, and Lord (2007) also conducted a factor analysis of children's performance on the ADOS. The authors found that a two-factor model of social affect and restricted, repetitive behaviors provided the best fit based on the data analyzed. This was contrary to the original scoring recommendations for the ADOS, which consisted of separate communication and reciprocal social interaction domains. A restricted, repetitive behaviors domain was also included in the original scoring; however, the ratings from this domain did not contribute to the final score. The authors recommended changing the scoring algorithm for the ADOS to reflect their findings.

Follow-up studies (Gotham et al., 2008; Oosterling et al., 2010) examined administration of the ADOS using a two-factor scoring algorithm (social affect and restricted, repetitive behaviors) as opposed to the traditional scoring method. Diagnostic validity of the ADOS was found to be higher using the revised algorithms than with the originally recommended algorithms.

Consequently, the second edition of the ADOS (Lord, et al., 2012) revised the scoring algorithm to include a single communication and socialization domain, termed social affect. The second domain was restricted and repetitive behavior. In the ADOS-2, restricted and repetitive behaviors are part of the scoring algorithm, though they were not included in the original ADOS scoring recommendations.

Analysis of the Autism Spectrum Rating Scales (ASRS)

Exploratory factor analysis was conducted on the questions from the ASRS (Goldstein & Naglieri, 2010) as part of the evaluation instrument's validation process. For children aged 2-5 years, results yielded two factors, rather than the traditional three. The authors agreed with previous research and reported that instead of being separate factors, social interaction and communication fit best as one combined social communication domain. The second domain identified was unusual behaviors. For children aged 6-18 years a third factor of self-regulation was also identified.

Summary of Analyses

While some of the reviewed studies found a single factor based on the tests examined, and others identified multiple factors, one constant theme among the research emerged -- there was not sufficient evidence to support separate domains for communication and social impairments when evaluating for autism. Rather, it was recommended that these areas be combined into a single social communication domain. Research also supported the inclusion of a second domain addressing repetitive and stereotyped behaviors. Accordingly, the *DSM-5* recommends a two-factor model for autism spectrum disorder consisting of deficits in social communication and the presence of "restricted, repetitive patterns of behavior" (APA, 2013, p. 50).

Autism in Preschool-Age Students

To understand how autism is expressed in preschool-age students, it is important to understand typical development for children at this age. The following section will address preschool development and early signs of autism.

Social and Communication Development in Preschool-Age Students

Between the ages of 2 and 5, typically developing children experience an explosion in the development of communication and social skills. Per the Mayo Clinic's (2012) established developmental milestones a typical 2-year-old will usually have about 50 words, which increases to 500 or more by age 3. By the time children are 4, they are expected to answer simple questions and have progressed from combining just two words to speaking in complex sentences. The development of adjectives and pronouns occurs from ages 2-3, and prepositions and future tense language emerges at ages 4 and 5. In addition, according to the Pennsylvania Department of Education (PDE) and Department of Public Welfare (DPW) Early Learning Standards (2009), pre-kindergarten students should be able to engage in reciprocal conversations with others, verbally describe experiences, and incorporate nonverbal gestures into their communication.

Socially, young children begin imitating caregivers and other children in play schemes before age 3 (Mayo Clinic, 2012). Turn taking with others and expressing affection appropriately are also expected at age 3. Cooperation with other children during play and a desire to be like their peers emerges at 4 and 5 years of age. Most children develop the concept of a best friend by age 4. Play skills also begin to develop by age 2 (Carter, Davis, Klin, & Volkmar, 2005). Children engage first in functional play (using objects as intended) and then progress to pretend or symbolic play (engaging with objects beyond their physical properties). PDE and DPW Early Learning Standards (2009) indicate that when interacting with adults pre-kindergarten students should be able to express their emotions in a socially appropriate manner, label their feelings accurately, and ask for help when needed. Pre-kindergarten students should also be able to engage in play with peers for a sustained period of time, and begin to resolve conflicts with other children.

Characteristics of Autism Identifiable in Early Childhood

For some children, this pattern of development does not emerge as expected, which may raise concerns regarding a possible developmental delay or an autism spectrum disorder. The earliest signs of autism in early childhood are often what are called negative signs, or the lack of appropriately developing communication and social skills as described above (Filipek et al., 1999). Negative signs, while easy to detect, can be more difficult to attribute to autism than what are known positive signs (Coonrod & Stone, 2005). Positive signs include behaviors typically associated with autism such as "the presence of unusual sensory and motor behaviors" (Coonrod & Stone, p. 708). The presence of both positive and negative signs is what helps differentiate the diagnosis of autism from other developmental delays.

These signs can be observed early in life, and parents often report noting differences in development as early as 6-12 months of age (Carter et al., 2005). Even Kanner (1943) noted that some characteristics of autism could be observed at infancy, and reported that at a very young age the children in his case studies did not alter their position to show anticipation of being picked up by their caregivers. Carter et al. (2005) explained that children with autism show early social deficits, or negative signs including lack of appropriate eye gaze and eye contact, lack of interest in social speech, lack of joint attention (calling another person's attention by pointing and looking at an object), lack of imitation, lack of appropriate functional and symbolic play, and lack of peer relations. Positive signs, such as engaging in repetitive or stereotyped actions with toys and other objects, could be observed by age 2.

Retrospective video analysis has become an area of growing research, in which videos of milestones such as first birthday parties are reviewed by researchers after a child has been diagnosed with autism, to determine what characteristics were observable prior to the diagnosis. In a review of this research, Crane and Winsler (2008) found the following negative signs, all observable prior to 24 months of age: Lack of shared attention with the mother, low frequency of social interaction, lack of symbolic play, lack of response to name, lack of pointing to reference objects, failure to seek contact with others, and failure to look others in the face (p. 248). Positive signs included excessive mouthing of objects, paradoxical reaction to sounds, and excessive exploratory/sensory activity with objects. Other video analysis studies found that deficits in pointing, showing objects, social smiling, and response to name could be observed as early as 9-12 months of age (Coonrod & Stone, 2005). Overall, researchers agree that difficulties with eye contact, showing objects to others, response to name, pretend play, and imitation noted by 18-24 months are highly correlated with a later diagnosis of autism (Baron-Cohen et al., 2000; Filipek et al., 2000; Robertson et al., 1999; Robins et al., 2001).

Preschool-age children with autism show similar deficits, including difficulties with imitation, functional play, sharing and responding to affective information, and engaging in joint attention (Coonrod & Stone, 2005). Other negative signs noted at the preschool age include difficulties with expressing emotion, as well as a general lack of prosocial behaviors such as "sharing, helping, offering comfort, offering affection, greeting others, and responding to humor" (Robertson et al., 1999, para. 4). Positive signs at this age typically include the more classic and well-known characteristics of autism such as hand flapping, aligning toys or objects, repetitive patterns of play, sensitivities to sensory input, and restricted areas of interest.

Autism Identification

While emerging research suggests that there may be genetic predisposition to autism spectrum disorders (Rutter, 2005), at this time there is no conclusive genetic or medical test capable of diagnosis. Currently, the diagnosis of autism is made with clinical judgment based on observed characteristics of autism (Levy et al., 2007). This can be done by matching observed and reported characteristics with *DSM-5* criteria for autism (APA, 2013; Schwartz & Davis, 2008), or through the use of autism-specific evaluations instruments, as described below.

Best Practices in Autism Identification

In Best Practices in School Psychology IV, Ikeda (2002) presented the record review, interview, observation, and testing (RIOT) model as best practice for autism identification in schools. Record reviews included examination of both school and clinical records to determine the severity and pervasiveness of behaviors. Interviews of caregivers can be both structured, such as the ADI-R (Rutter, Le Couteur, & Lord, 2003), and unstructured, in order to gather a basic developmental history. Observations should be conducted in multiple settings to provide an idea of students’ social interactions and repetitive behaviors. Ikeda (2002) included diagnostic assessments such as Childhood Autism Rating Scale (CARS, since updated to the CARS-2; Schopler, Van Bourgondien, Wellman, & Love, 2010) and the ADOS (since updated to the ADOS-2; Lord et al., 2012) as part of the observational portion of the evaluation. Testing included standardized measures of ability, achievement, and adaptive behavior, as needed to ascertain present levels of functioning.

This method of autism assessment would be useful if the evaluator were fairly certain of a likely autism diagnosis for the child; however, significant amounts of time could be spent completing assessments and evaluations that may or may not be needed. It is also lacking in specifics as to when and how practitioners would arrive at a diagnosis of autism. The Pennsylvania DPW created the Pennsylvania Autism Assessment and Diagnosis Expert Work Group in order to generate a comprehensive set of recommendations for autism diagnosis in the state of Pennsylvania, which proposed a three-stage model of autism identification (Levy et al., 2007). The workgroup "included parents of children with autism and professionals in multiple disciplines and subspecialties including audiology, epidemiology, neurology, nursing, occupational therapy, psychiatry, psychology, social work, special education and speech/language pathology" (p. 8).

In this model Levy et al. (2007) recommended that when concerns are noted children should receive an evaluation of development through either the educational or behavioral health system. Members of the evaluation team may include (as needed) parents, psychologists, speech-language pathologists, occupational therapists, clinical social workers, behavioral analysts, developmental pediatricians, child psychiatrists, and special education teachers. Overall, three stages were recommended as part of the diagnostic process. Stage 1 included gathering background information, reviewing appropriate records, and administering caregiver-completed questionnaires about the child's developmental history. If the findings were consistent with autism, practitioners were advised to move to Stage 2.

At Stage 2 a comprehensive developmental evaluation was recommended, based on the needs identified in Stage 1. Stage 2 also included observations and autism-specific assessments. Autism-specific evaluations that may be used included observational tests and parent/teacher completed checklists. Per the workgroup's recommendations, a diagnosis of autism could be made at Stage 2 if all data were confirmatory and there were no other questions. If there were additional questions or more information was needed, practitioners were advised to move to Stage 3.

During Stage 3 it was recommended that highly qualified clinicians review the data gathered during Stages 1 and 2 and conduct specialized diagnostic evaluations for autism. Assessments recommended at this stage included the

ADOS (since updated to the ADOS-2) and/or the ADI-R. Levy et al. (2007) noted that not every program or facility would have the ability to complete a Stage 3 evaluation, and should have a referral system in place so that children may receive these evaluations elsewhere if needed. This model has the advantage of a stepwise progression of autism identification, where decisions to proceed to additional testing are made throughout the evaluation process.

Autism Evaluation Instruments

Various instruments have been developed for use in autism identification. Commonly used rating scales include the CARS-2 (Schopler et al., 2010), the Gilliam Autism Rating Scale, Second Edition (GARS-2; Gilliam, 2004), the Asperger Syndrome Diagnostic Scale (ASDS; Myles, Jones-Bock, & Simpson, 2001), and the ASRS (Goldstein & Naglieri, 2010).

Follow-up studies of the GARS/GARS-2 and the ASDS have found questionable validity due to low sensitivity rates, and these measures are generally not recommended for diagnostic purposes (Norris & Lecavalier, 2010; South et al., 2002). In addition, the ASDS is only standardized for children aged 5 and older, rendering it problematic for use in early intervention (EI) programs. The CARS-2 and the ASRS, however, show promise as valid and reliable tools for assisting with an autism diagnosis (Goldstein & Naglieri, 2011; Schopler et al., 2010). Also of note, the ASRS was designed with a two-factor model of autism identification, closely resembling the APA's (2013) new recommendations for the diagnosis of autism spectrum disorder, whereas other older instruments were designed based on *DSM-IV* and *DSM-IV-TR* criteria (APA, 1994, 2000). The CARS-2 and the ASRS are suitable instruments for use as part of Stage 2 autism-specific testing in the Pennsylvania autism diagnosis model described above (Levy et al., 2007).

Additionally, certain diagnostic instruments, specifically the pairing of the ADOS/ADOS-2 (Lord et al., 1999, 2012) and the ADI-R (Rutter et al., 2003), have come to be regarded as the "gold standard" in autism identification (Akshoomoff, Corsello, & Schmidt, 2007; Ikeda, 2002; Levy et al., 2007; Lord et al., 2000; Oosterling et al., 2010). As discussed previously, the new diagnostic algorithm for the ADOS-2 is based on a two-factor model of autism, in agreement with the new *DSM-5* recommendations (APA, 2013). These instruments are suitable for use as part of Stage 3 diagnostic autism testing in the Pennsylvania autism diagnosis model (Levy et al.).

The ADI-R (Rutter et al., 2003), is an extended interview conducted with a caregiver designed to elicit a full range of information needed to produce a diagnosis of autism. The interview questions address a child's background, early development, behavior, language acquisition, social development, repetitive and stereotypical behaviors, and current functioning.

The ADOS-2 (Lord et al., 2012) is a direct observation of the essential deficits of autism. While the ADOS-2 is a self-described semi-structured observation, of all the instruments described in this section, it alone involves direct interaction with the child being evaluated. Play situations are arranged to elicit behaviors relating to communication, social skills, and stereotyped mannerisms to allow professionals to observe the child's responses. One characteristic of the ADOS-2 in particular makes it excellent for use in the schools. It has varying modules that can be used for children with no language, children with phrase speech only, and children with highly developed verbal skills.

Understanding the recommended models for autism identification and having a basic familiarity with evaluation instruments, however, is not enough for a school psychologist, or any practitioner, to identify a child with an autism spectrum disorder. Experts agree that the diagnosis of autism should be made only by those who have specialized training, significant clinical experience, and extensive expertise with autism (Levy et al., 2007; Lord & Corsello, 2005; Schwartz & Davis, 2008).

Considerations for Autism Identification in Young Children

Studies have demonstrated that autism can be reliably and validly diagnosed in children as young as 2 years of age (Baird et al., 2000; Pandey et al., 2008; Robins et al., 2001; Stone, Coonrod, & Ousley, 2000). It is noted that parents of very young children, however, may have difficulty answering questions about autism-specific characteristics. Specifically, parents may have difficulty distinguishing between their child's ability to follow along with known, familiar social routines and their ability to engage in "spontaneous, socially motivated interactions" (Lord & Corsello, 2005, p. 730). Expert knowledge in this area is needed to obtain accurate parent report for children at young ages.

Another crucial consideration when evaluating young children for autism is in choosing instruments that are appropriate to the age that is being assessed (Lord & Corsello, 2005). Instruments appropriate for use with very young children include the CARS-2; the ADOS-2, which includes a toddler module along with other modules appropriate for children with limited language; or the ASRS, which has a specific preschool-age form. As mentioned above, a thorough understanding of typical development of young children, along with an understanding of the positive and negative signs of autism that are likely to be present in young children, is necessary for conducting evaluations with preschool-age children. Again, the diagnosis of autism, regardless of the child's age, should be made only by those who have specialized training, significant clinical experience, and extensive expertise with autism (Levy et al., 2007; Lord & Corsello, 2005; Schwartz & Davis, 2008).

Necessity for Accurate Autism Screening

The best autism assessment model and the most thorough understanding of preschool-age development are meaningless if children are not accurately screened to ensure that they receive these evaluations when necessary. As mentioned previously, in the state of Pennsylvania a school psychologist's involvement is not legally mandated in EI evaluations for children suspected of having developmental delays only (Chapter 14 Special Education Regulations, 2008, §14.153). Accurate screening is a necessity to ensure that children receive appropriate evaluations when autism is a possible concern.

Screening for autism in EI programs is vital, as "children who have been identified as having delays but for whom the specific diagnosis of autism has not yet been made may not receive the types of specialized intervention services found to be the most effective with this population" (Rogers, as cited in Coonrod & Stone, 2005, p. 709). Research overwhelmingly supports the idea that early services and interventions result in improved outcomes for children with autism (Crane & Winsler, 2008; Levy et al., 2007; National Research Council, 2001; Robins & Dumont-Mathieu, 2006).

Screening Measures for Autism

Screening instruments are best understood when divided into two levels: Level I and Level II screening instruments (Coonrod & Stone, 2005). Level I screeners are instruments designed "to identify children at risk in the general population" (Robins & Dumont-Mathieu, 2006, p. S112). Level II screeners, in contrast, are to be administered to a "selected group of children already considered to be at increased risk" for developmental disabilities (Robins & Dumont-Mathieu, p. S112; see also Norris & Lecavalier, 2010). Level II screening measures will be discussed first, as they are appropriate for use with the high-risk preschool-age students who have been referred to EI programs.

Level II Screening Measures

Level II autism screeners are intended to distinguish children with autism from children with other developmental delays (Coonrod & Stone, 2005; Robins & Dumont-Mathieu, 2006). The following is a discussion of several well-known Level II autism screening instruments.

The ASRS (mentioned previously) has a short form in addition to the standard form (Goldstein & Naglieri, 2010). This 15-item Likert scale short form is meant to be used as a screening tool to differentiate students with autism from those without. This tool can be used for children as young as 2 years of age. This screening measure demonstrates excellent discriminative validity, with high sensitivity and specificity (.94 and .92 respectively; Wilkinson, 2011).

The SRS-2 (Constantino, 2012) is another Level II autism screening tool. Until its update in 2012, the SRS was only standardized for children aged 4 and older, prohibiting use of this tool with a large portion of preschool-age students (Constantino & Gruber, 2005). The updated SRS-2, however, added a preschool age form that can be used for children 2.5 to 4.5 years of age. The preschool age form is relatively new, but sensitivity and specificity have been estimated at .85 and .75, respectively (Wilkinson, 2011).

The Social Communication Questionnaire (SCQ; Rutter, Bailey, & Lord, 2003) is a 40-item yes/no questionnaire that provides a cutoff score of 15 or more failed items as the referral for autism evaluation. While significant research has been conducted on the SCQ (Norris & Lecavalier, 2010; Wilkinson, 2011), it is only standardized for use with children above 4 years of age; consequently a large portion of preschool-age children could not be screened with this measure. In addition, Corsello et al. (2007) found that the SCQ was significantly less sensitive when used to screen children under the age of 7. For children aged 4-7 they recommended a lower cutoff score of 11, as many young children with autism passed the screener using the recommended cutoff score of 15.

The Pervasive Developmental Disorders Screening Test-II (PDDST-II; Siegel, 2004) is an autism screening tool for children aged 12-48 months. The PPDST-II is broken into three stages. Stage 1 is for screening at a pediatrician's office (or similar), where children who failed the screener would be referred for further evaluation. Stage 1 is best defined as a Level I screener and is discussed in the following section. Stages 2 and 3 most closely align with Level II screening parameters. Stage 2 is for use in developmental clinics, with sensitivity and specificity reported at .73 and .49, respectively. This stage is meant to distinguish children who are possibly in need of referral for autism testing. Stage 3 is for use in autism-specific clinics, with reported sensitivity and specificity of .58 and .60, respectively. Overall, Stages 2 and 3 of the PPDST-II had disappointing sensitivity and specificity as compared to Stage 1 (Dumont-Mathieu & Fein, 2005).

EI programs are in need of reliable and valid screening tools that are cost-effective, and can be administered to all children to make accurate referrals for autism testing. Based on the description of Level II screeners (i.e., screening instruments meant to "differentiate children at risk for autism from those at risk for other developmental disorders," [Coonrod & Stone, 2005, p. 709]), it would seem that these types of screeners would be ideal for use in EI programs. Level II screeners can certainly be used once there is already a suspicion of autism to help make diagnostic and evaluative decisions (as part of Stages 1 or 2 in the autism identification model described previously; Levy et al., 2007), but they cannot be administered to every child who referred to an EI program. The Level II screening tools described above must be purchased on a per-protocol basis, which is cost-prohibitive for use with every referred child in many EI programs. In addition, the SCQ has lower sensitivity when being used for children under 7, and is not standardized for children under 4, which limits its practical application for EI programs. Further, these measures require significant amounts of time to administer, score, and interpret (Coonrod & Stone, 2005), making their use for all children referred to EI programs problematic.

Level I Screening Measures

Level I autism screening instruments are meant to distinguish children at risk for autism in the general population (Coonrod & Stone, 2005; Robins & Dumont-Mathieu, 2006). They should be brief to administer as well as cost-efficient to support their use with large numbers of children (Coonrod & Stone, 2005; Levy et al., 2007).

The PDDST-II (mentioned above; Siegel, 2004) is an autism screening tool for children aged 12-48 months, which is divided into three stages. Stage 1 most closely aligns with the parameters of a Level I screening instrument. Stage 1 is used for screening children at a pediatrician office visit or similar situations. Children who fail the screener are referred for further evaluation. Sensitivity for Stage 1 was reported at .91-.92. Unfortunately, while this measure show promising statistical strength, each protocol used must be purchased, which limits the practical application in an EI program where every referred child needs to be screened.

Various other low-cost Level I screeners have been developed for use in other countries, but have not been validated in the United States, such as the Developmental Behaviour Checklist— Early Screen (Gray & Tonge, as cited in Allison et al., 2008) and the CHecklist for Autism in Toddlers (CHAT; Baron-Cohen, Allen, & Gillberg, 1992). Still other Level I screening measures address populations much younger than the 3-5 year age range of EI programs, including the Communication and Symbolic Behavior Scales Developmental Profile (Wetherby, Allen, Cleary, Kublin, & Goldstein, 2002), the Screening Test for Autism in Two Year Olds (STAT; Stone, Coonrod, & Ousley, 2000), the Systematic Observation of Red Flags for Autism Spectrum Disorders in Young Children (Wetherby & Woods, 2004), and the Early Screening of Autistic Traits Questionnaire (ESAT; Swinkels, as cited in Allison et al., 2008). These screeners are typically used for children aged 12-24 months.

There seems to be a gap in the available screeners for the preschool-age population. The available Level II screeners tend to focus on children ages 4 and older, and additionally could be cost-prohibitive if they were to be administered to all children referred for EI services. Similar problems are noted with Level I screeners. Of the Level I screeners that had the advantage of being free, they tended to focus on very young children, often as young as 12-24 months. There seems to be a lack of widely available screening instruments that can be universally administered to preschool-age students who are referred to EI services in order to make accurate referrals for autism evaluations. In essence, what is needed for EI programs is an autism screener with the ease, utility, and cost-effectiveness of a Level I screener that has been validated for use with high risk, or Level II, populations.

In reference to the Level I screeners described above, Allison et al. (2008) concluded that "while many of these instruments have been tested on referred populations and have good psychometric properties, none have been evaluated in the general population" (p. 1415). Similarly, Mawle and Griffiths (2006) found in a review of studies examining Level I autism screening instruments in children under 4, only studies examining the CHAT (Baron-Cohen et al., 1992) and the M-CHAT (Robins et al., 1999), which was based on the CHAT, met criteria to be included in their review. All other studies reviewing universal screening measures for autism did not meet rigorous requirements for inclusion because of lack of appropriate sample sizes, lack of information regarding sensitivity and specificity, or the lack of simplistic scoring and administration procedures to allow for ease of use. The CHAT and M-CHAT are discussed in detail below.

Background and Development of the M-CHAT

The M-CHAT is a universal, or Level I, screening instrument that was initially developed for very young children (16-30 months; Robins et al., 2001), although some follow-up research has been conducted to examine its use with older children. It has also been studied in high-risk populations. The development, validation studies, and follow up research for the M-CHAT are discussed in this section.

Development of the CHAT. The original CHAT was designed in Great Britain and was completed through a combination of nine parent interview questions and five observation items (Baron-Cohen et al., 1992). The CHAT was developed as a universal screening measure for children 18 months of age, with the observation portion completed in the child's home by a home health visitor. Initial studies were very promising, with the authors reporting sensitivity and specificity rates as high as 100% (Baron-Cohen et al., 1992, 1996). Follow-up studies; however, found extremely low sensitivity rates (Baird et al., 2000; Baron-Cohen et al., 2000; Mawle & Griffiths, 2006; Robins et al., 2001) and reported that the majority of participants who were later diagnosed as having autism had not been recognized when the CHAT was first administered (Eaves, Wingert, & Ho, 2006). Other problems noted with the CHAT were that the observation portion was designed to be completed in part by a home health visitor (which does not exist in the United States), the short observation period was not enough time to observe all possible characteristics of autism, and the screener was only designed to identify severe cases of autism, rather than all autism spectrum disorders (Robins et al., 2001).

Development of the M-CHAT. The M-CHAT was developed by Diana Robins as part of her doctoral dissertation through the University of Connecticut (Robins et al., 2001). The M-CHAT pulls the nine parent interview items from the original CHAT, but excludes the observation portion. In addition to the items from the CHAT, 21 new questions were developed as part of the initial study, for a total of 30 questions. The initial sample included 1122 children (either 18 or 24 months) screened through family doctors/pediatricians and 171 children (average age 26 months) identified through EI programs, for a total of 1293 participants. After analysis of participant responses, the total number of questions was reduced to 23. Each item is answered by the parent as either Yes or No, and is then scored as Pass or Fail (see Appendices A and B).

In the initial validation study, Robins et al. (2001) found that the M-CHAT's sensitivity (participants who received an autism diagnosis and failed the M-CHAT) was .87-.97. Its specificity (participants who did not receive an autism diagnosis and passed the M-CHAT) was .95-.99. The M-CHAT had a positive predictive value (PPV) of .36 (true positives divided by all screened positive cases) based on scoring guidelines of failing any two critical items or any three items. When a follow-up interview was incorporated the PPV rose to .68. Critical items were identified as children's ability to take an interest in other children, use their index finger to point to objects of interest, bring objects to their parents to show them something, imitate facial expressions, respond to their name, and follow a parent's point in order to look at something across the room.

Per the M-CHAT's instructions for use, the authors determined that there would be a high false positive rate for the M-CHAT (participants who failed the M-CHAT but did not receive an autism diagnosis), to ensure that the maximum number of participants who truly did have autism also failed the screener (Robins et al., 1999, 2001). This was acceptable as participants who failed the M-CHAT were likely in need of further evaluation to determine the presence of developmental delays, even if they did not meet criteria for autism (Robins, 2008; Robins et al., 1999, 2001).

Follow-up research studies. In 2008 a replication study was completed with 3793 participants between the ages of 16-30 months, yielding similar results to the original findings (Kleinman et al.). An overall PPV of .36 was reported when using established scoring guidelines, which rose to .74 when a follow-up interview was conducted. In this study, the authors also calculated separate PPVs based on whether participants were from the low-risk general population, or from a high-risk population already identified as having concerns. The low-risk sample yielded a PPV of .11, increased to .65 when follow-up interviews were included. The high-risk sample yielded a PPV of .60, which increased to .76 when follow-up interviews were included. This would suggest that the M-CHAT is a useful tool for autism screening in students when concerns have already been noted. The authors also completed follow-up interviews two years later with 15 participants about whom concerns had been noted, identified as "possible missed cases" (p. 835), in order to determine if any children later received an autism diagnosis after passing the M-CHAT. Of these 15 cases, the authors found that seven had been diagnosed with an autism spectrum disorder, one was diagnosed with developmental delays, three had language delays, and four were not identified as having any disability.

Robins (2008) completed an additional study investigating 4797 children between the ages of 16-30 months, who were screened with the M-CHAT at pediatrician office visits. Overall, 466 children failed the M-CHAT, which was decreased to 61 after follow-up interviews were completed. Of these, 41 children received evaluations and 21 were diagnosed with an autism spectrum disorder. An additional 17 were reported to have developmental delays unrelated to autism, and three had no delays. The PPV for children who failed the M-CHAT was only .058; however, the PPV increased dramatically when follow-up interviews were completed, to .57.

In reviewing these studies, it appears that the PPV of the M-CHAT is relatively low when it is used as a stand-alone screening measure. In every study, however, the PPV increased substantially when the M-CHAT follow-up interview was completed with parents. This is not as easy a solution as it may seem, however, as universal autism screening measures need to be efficient and easy to administer. "In an ideal situation they could be given without taking professional time and quickly scored, flagging children who need further investigation" (Eaves et al., 2006, p. 230). Despite the evidence that the follow-up interview increases the predictive value of the M-CHAT, many sites that routinely use the M-CHAT do not complete the follow-up interview as it decreases efficiency and was only recently made publicly available.

M-CHAT research examining subpopulations. Although the M-CHAT was designed to be used as a universal screener given at pediatrician office visits, it has been used and researched in other settings and with children older than 16-30 months. Eaves et al. (2006) studied the use of the M-CHAT at an autism clinic in Canada, examining its use with 84 children 2 and 3 years old (with a mean age of 37 months) who had been referred for autism evaluations. The follow-up interview was not used as part of this study.

Overall, it was found that the sensitivity of the M-CHAT was relatively high in this sample (.77-.92), but specificity was much lower than had been previously reported (.27-.43). In addition, there were a substantial number of false negatives reported, or children who received an autism diagnosis from the clinic but had not been identified by the M-CHAT. These results are in contrast to the findings of previous M-CHAT studies, and may have been due in part to the M-CHAT's use with older children.

Additional studies have been completed to gain information about the M-CHAT for children older than 30 months and for children who have already received diagnoses. Snow and Lecavalier (2008) examined the M-CHAT in a sample of preschool age students in a clinical setting (aged 18-70 months, n = 84). In this sample 54 children had received an autism spectrum disorder diagnosis and the remaining 28 children had other developmental delays that were not specific to autism. The authors found that the cutoff criteria of failing any three items on the M-CHAT identified 77% of children with autism; however, the cutoff criterion of failing two critical items was less accurate. Sensitivity and specificity rates overall were lower in this sample (.70-.88 and .38, respectively) than what was reported by the initial M-CHAT validation research (Robins et al., 2001). The authors noted that the comparison group consisted of children identified as having developmental delays, which may have decreased the accuracy of the M-CHAT, "because the behavioral profiles of children with PDDs are more similar to children with [developmental delays] than typically developing children" (Snow & Lecavalier, p. 640). This research is vital in determining whether the M-CHAT is capable of distinguishing autism from other delays and disorders present in preschool-age students.

Overall, the studies examining the M-CHAT's use with children who are older than the intended age range, and who are high risk (children who have already been diagnosed with delays) seem to indicate lower sensitivity rates and lower PPV (Eaves et al., 2006; Snow & Lecavalier, 2008) than the initial validation studies reported (Kleinman et al., 2008; Pandey et al., 2008; Robins, 2008; Robins et al., 2001). It was also noted that a follow-up interview was required in all the validation studies to boost the PPV of the M-CHAT to acceptable levels, as otherwise significant amounts of false positives were identified. This follow-up interview was only recently made publicly available; therefore, it was not included in many of the M-

CHAT follow-up studies, nor is it included in the present research study.

Of note, at the 2010 International Meeting for Autism Research, Robins et al. presented preliminary findings regarding a new scoring algorithm for the M-CHAT, designed to decrease the large amount of false positives previously reported. These scoring criteria, known as the Best7, identified seven critical items rather than the previous six (see Appendix B). The recommendation of referral for an autism evaluation based on failure of two critical items or any three items remained the same. The authors reported that this change improved the PPV of the M-CHAT from .04 to .18 without the follow-up interview. The PPV improved from .52 to .61 with the use of the follow-up interview. The results of this research have not yet appeared in the published literature, and the scoring instructions on the M-CHAT website (www.mchatscreen.com) continue to recommend the six critical items scoring approach.

Rationale for Study

Autism was originally thought be a rare disorder, with initial prevalence rates estimated at around 4-5 in 10,000 children (or approximately 1 in 2000) based on the earliest epidemiological studies conducted in the 1960s in the United Kingdom (Lotter, 1966). In the 1990s autism in the

United States was estimated at anywhere from 1 in 500 to 1 in 1000 individuals, with reports varying significantly (Filipek et al., 1999, 2000; Frombonne, 2005). As of 2007 autism was being reported to occur at a rate of 1 in 150 children (Layne, 2007). In 2012 it was reported at a rate of 1 in every 88 children in the United States, and even more alarming, 1 in 54 boys (Baio). In its most recent update, the Centers for Disease Control and Prevention (CDC) estimated that autism now occurs in 1 of every 50 school-aged children (Blumberg et al., 2013).

Opinions vary on whether the prevalence of autism is truly increasing, or whether the increase in reported cases is actually due to better diagnostic accuracy and expanded definitions of autism spectrum disorders (Frombonne, 2005; Wing & Potter, 2009). Regardless of whether autism rates are actually increasing or whether diagnostic methods and understanding of autism has improved, one thing is clear -- it cannot be claimed that autism is a rare disorder anymore, with nearly 2% of the school-age population identified as having this disorder (Blumberg et al., 2013).

With this increase, school psychologists are increasingly called upon to identify students with autism and recommend appropriate services in the schools (Wilkinson, 2010, 2011). Despite the fact that the majority of children with autism are first identified by the school systems (Glascoe, 2000; Palfrey, Singer, Walker, & Butler, 1987; Yeargin-Allsopp et al., 2003), there is often a significant delay in diagnosis from the time that concerns are first noted (Crane & Winsler, 2008; Levy et al., 2007; Wilkinson, 2011). Early identification is vital, as research has shown that when interventions are provided at an early age outcomes for children can be improved (Crane & Winsler, 2008; Levy et al., 2007; National Research Council, 2001; Robins & Dumont-Mathieu, 2006). However, without effective early screening, children will not have access to needed EI services that are autism-specific (Crane & Winsler, 2008).

While screening is necessary to ensure that children receive autism evaluations, there are several problems inherent with the use of the M-CHAT as a screening instrument in EI programs. First, the instrument is used above the age range for which it was intended. While there is some research examining its use with older children (Eaves et al., 2006; Snow & Lecavalier, 2008; Yama, Freeman, Graves, Yuan, & Campbell, 2012), far more is needed before the M-CHAT can be reliably recommended for use with preschool-age students (Robins, 2008). Second, the majority of research focuses on the M-CHAT's use in clinical or university settings, not in schools (Dumont-Mathieu & Fein, 2005; Kleinman et al., 2008; Mawle & Griffiths, 2006; Pandey et al., 2008; Robins, 2008). Third, the M-CHAT was originally intended to be a universal, or Level I, screener (Robins & Dumont-Mathieu, 2006). The vast majority of the M-CHAT's initial validation sample participants were Level I (n = 1122), with only 171 participants identified as high-risk (Robins et al., 2001), with similar ratios reported in follow up studies (Kleinman et al., 2008). While replication studies indicated higher PPV for the M-CHAT in high risk populations than in the general population (Kleinman et al., 2008), other studies have reported lower specificity, sensitivity, PPV and NPV for the M-CHAT when examined exclusively with high-risk children (Eaves et al., 2006; Snow & Lecavalier, 2008). Because the M-CHAT was intended to be given to all children, it was designed with a high false positive rate, meaning that more children than just those with autism would fail the screener. In addition to ensuring that all children in need of autism evaluations would receive accurate referrals, the authors reasoned that any child who failed the M-CHAT was almost certainly in need of a developmental evaluation, regardless of whether he or she ultimately received an autism diagnosis (Robins, 2008; Robins et al., 1999, 2001). As a Level I screener, this rationale had excellent practical purposes. On the other hand, when the M-CHAT is used in EI programs, where it is already known that every child who is referred to the program is in need of a developmental evaluation, a Level I screener is administered to a population already known to be at risk. This begs the question of whether the use of the M-CHAT with this population will lead to an even higher rate of false positives than was originally intended. If the false positive rate is too high, the ability of the M-CHAT to accurately identify students in need of autism evaluations in EI programs is significantly diminished. Further research is needed to examine the M-CHAT's utility as a Level II screening instrument, specifically in educational settings.

Despite these concerns, the M-CHAT continues to be used for preschool-age students primarily due to its ease of use and cost-efficiency, in addition to the lack of other appropriate measures to be used in its place. This much needed research intends to examine the M-CHAT's use in a preschool-age special education sample in order to determine the differences in performance for students with autism and students with other developmental delays.

Summary

This chapter provided an overview of the history of autism and the current changes in the understanding and definitions of autism. The recommendations for diagnosing autism have changed significantly since it was first proposed as a diagnostic category in 1943. Characteristics of autism are observable in preschool-age students, indicating that early diagnosis is appropriate and necessary to allow students to receive EI services. Research relating to autism screening tools for preschool-age children indicates that more evidence is needed to support the use of the M-CHAT with this population. The purpose of this study is to examine the utility of the M-CHAT for preschool-age special education students. This study will also determine which items on the M-CHAT are most associated with students' receiving an educational classification of autism, and will additionally attempt to determine whether a two-factor model of autism identification (i.e., social communication deficits and unusual/repetitive behaviors) or a different model best describes parent ratings on the M-CHAT.

CHAPTER III

METHOD AND PROCEDURES

The purpose of this study was to examine the utility of the Modified Checklist for Autism in Toddlers (M-CHAT; Robins, Fein, & Barton, 1999) in a preschool-age special education sample. This study also examined which items on the M-CHAT were most associated with students' receiving an educational classification of autism, as discriminated from students with other developmental delays. In addition, this study attempted to determine whether a two-factor model of autism identification (i.e., social communication deficits and unusual/repetitive behaviors) or a different model best describes parent ratings on the M-CHAT.

Design

As there can be no random assignment of groups, the design of this study is a retroactive casual comparative study. No assignment method was used in this study, as the focus is analyzing preexisting data. Two groups were included in this study: Students with autism and students with other developmental delays, as per their educational classification.

Population

This research study occurred at the Berks County Intermediate Unit (BCIU), in Reading, Pennsylvania. Berks

County is located in Eastern/Central Pennsylvania and represents an extremely diverse population. Reading was rated as being the nation's poorest city until September of 2012; it is currently ranked the sixth poorest with a poverty rate of 40.1% (Brudereck, 2012). There is a large Latino population in the city of Reading, where 58% of residents identify themselves as such. There is a sharp divide between the average household income in Reading ($34,083) and Berks County as a whole ($66,641), which includes inner-city Reading, suburban areas, and rural farming communities.

The BCIU provides Early Intervention (EI) services for the 18 school districts in Berks County. EI services are special education services for students aged 3-5, identified as having a disability. Approximately 1800 preschool-age students were being served by the BCIU EI program at the time of this study. All students were at least 3 years of age at the time the archival data were gathered; however, it is noted that many students were 2 at the time the M-CHAT was completed, due to legal requirements to evaluate children who have already been identified as a toddler with a disability before their 3rd birthday to allow for an IEP to be implemented upon their turning 3 years of age (Individuals with Disabilities

Education Improvement Act [IDEIA], 2004, §303.209). Therefore, the age range of the population at the time the M-CHAT was completed was 30-72 months.

Sample

Inclusion Criteria

All students entered into the BCIU EI program's IEP Writer database who were identified as having an educational classification of autism (n = 133) as of May 2011 were included in this study. A random sample of students with other educational classifications who were enrolled in the program as of May 2011 was also selected to create a comparison group.

Systematic random sampling was used to obtain the comparison group. An alphabetical list of all students entered into the BCIU's IEP Writer database as of May 2011 (excluding students with autism and students for whom autism testing had been recommended) was generated by a BCIU EI employee volunteer. The total number of students on this list (n = 1064) was divided by the number of students in the autism group (n = 133). Based on this calculation, every eighth name on the list was selected for inclusion in the comparison group (n = 133). Thus, data were gathered for a total of 266 participants in this study.

Exclusion Criteria

Students for whom autism testing was recommended but not yet completed were excluded from this study. Prior to generating the comparison group, these students were removed from the list of students with other developmental delays. Students whose records were incomplete (i.e., M-CHAT not completed) were excluded from this study after archival records were gathered.

Participants

Of the 266 students initially selected to be included in this study, M-CHAT records were available for 222 students (n = 109 students with autism, n = 113 students with other developmental delays). A total of 44 students were excluded from data analysis due to missing M-CHAT forms. Table 1 describes the sex, age, and race of students whose records were included in the data analyses. It was noted that 30.2% of the participants were identified as Hispanic. In their study, Eaves, Wingert, and Ho, (2006) reported that the M-CHAT actually demonstrated slightly better sensitivity for classifying participants for whom English was a second language, as opposed to participants whose first language was English. It is unknown how many of the families in the current study spoke Spanish as their primary language; however, the M-CHAT is

Table 1

Sex, Age, and Race of Participants

	Autism Group		Comparison Group		Total	
	n	%	n	%	n	%
Sex						
Male	93	85.3	86	76.1	179	80.6
Female	16	14.7	27	23.9	43	19.4
Age						
2	64	58.7	56	49.6	120	54.1
3	30	27.5	39	34.5	69	31.1
4	14	12.8	16	14.2	30	13.5
5	1	.9	2	1.8	3	1.4
Race						
Caucasian	73	67.0	70	61.9	143	64.4
Hispanic	26	23.9	41	36.3	67	30.2
African American	9	8.3	1	.9	10	4.5
Other	1	.9	1	.9	2	.9
Total	109	100	113	100	222	100

provided to all parents with the English form on one side of the paper and the Spanish form on the other, allowing families to complete the screener in their preferred language.

Educational classifications of students in the comparison group included developmental delay, hearing impairment, other health impairment, orthopedic impairment, and speech or language impairment (Chapter 14 Special Education Regulations, 2008, §14.101; IDEIA, 2004, §300.8). Table 2 describes the educational classifications of students in the comparison group.

Table 2

Comparison Group Educational Classifications

Educational Classification	n	%
Developmental Delay	79	69.9
Speech or Language Impairment	26	23.0
Other Health Impairment	5	4.4
Orthopedic Impairment	2	1.8
Hearing Impairment	1	.9
Total	113	100

Measurement

Dependent Variable

The dependent variable in this study was parent ratings of student performance on the M-CHAT, an autism screening instrument. This included whether students passed or failed the M-CHAT, as well as responses to individual questions. The M-CHAT is administered to the parents of every child who is referred to the BCIU for EI services.

The M-CHAT is a 23-item, parent-completed checklist designed to identify young children (16-30 months) in need of evaluation for autism (Robins et al., 1999; see Appendix A). Each item is answered by the parent as either Yes or No, and is then scored as Pass or Fail. Items on the checklist are classified as critical or noncritical items. Based on discriminant function analysis, Robins, Fein, Barton, and Green (2001) recommended that children have failed the M-CHAT and should be referred for further evaluation for autism if they failed two or more critical items or any three items overall on the M-CHAT (see Appendix B). The initial validation study, which included 1293 participants, found that the M-CHAT yielded adequate internal reliability estimates of .85 for the entire checklist, and .83 for critical items using Cronbach's alpha (Robins et al., 2001).

Independent Variables

The independent variables in this study were participants' disability category (autism vs. other developmental delays) and the scoring method used (traditional vs. Best7). In 2010 Robins et al. presented preliminary findings regarding a new scoring algorithm for the M-CHAT, designed to decrease the large amount of false positives previously reported. These scoring criteria, known as the Best7, identified seven critical items rather than the previous six (see Appendix B). Overall interpretation of the M-CHAT remained the same; children have failed the M-CHAT and should be referred for further evaluation for autism if they failed two or more critical items or any three items on the M-CHAT.

Demographic information was obtained from the BCIU's IEP Writer database and from participants' main files. At the BCIU, students who are given an educational classification of autism through the EI program undergo at minimum an observation by at least one qualified professional (school psychologist, special education teacher, speech therapist, and/or occupational therapist), an in-depth developmental history gathered through parent interview, and diagnostic testing for autism. Most often, diagnostic testing includes the use of the Autism Diagnostic Observation Schedule (ADOS; Lord, Rutter, DiLavore, & Risi, 1999), which is considered the gold standard for autism identification (Lord et al., 2000). The ADOS is administered by a school psychologist and a speech therapist, with one person administering and one observing. Scoring is conducted by both professionals to ensure inter-rater reliability and agreement on final scores and diagnostic decisions. For a small number of students who do not meet minimum requirements for the use of the ADOS (developmental level below 18 months or mobility difficulties; Lord et al., 1999) diagnostic tests such as the Childhood Autism Rating Scale, Second Edition (Schopler, Van Bourgondien, Wellman, & Love, 2010) may be used. In addition to these diagnostic tests, other autism rating scales and screening instruments are used at the examiner's discretion to gather further data as needed. These include the Autism Spectrum Rating Scale (Goldstein & Naglieri, 2010), the Social Communication Questionnaire (Rutter, Bailey, & Lord, 2003), the Asperger Syndrome Diagnostic Scale, and the Gilliam Autism Rating Scale, Second Edition (Gilliam, 2004).

Bilingual interpreters and speech therapists are available as needed for students whose primary language is not English, to allow for federal regulations requiring evaluations to be administered in a child's native language whenever possible (IDEIA, 2004, §300.29). All parent-completed autism rating scales are available in both Spanish and English. Overton, Fielding, and Garcia de Alba (2007) reported historic under diagnosis of autism spectrum disorders in Hispanic populations, as well as a delayed age of diagnosis compared to Caucasian children. Specific concerns were that "misinterpretation ... may take place while observing the child when the examiner is not familiar with the culture" (Rhodes, Ochoa, & Ortiz, as cited in Overton et al., 2007, p. 1997), and that examiners may be hesitant to give this diagnosis to children of different cultures. The authors' recommendations for making diagnostic decisions with children of Hispanic descent did not differ greatly from recommendations discussed previously in Chapter 2 for evaluating any child suspected of having autism. Gold-standard diagnostic tests, including the ADOS and ADI-R were recommended, administered in the child and parents' native language, as necessary. Further, a thorough knowledge and understanding of the child's culture and heritage, along with an understanding of bilingual language development, are essential. In the city of Reading, cultural influences include not only a high percentage of Hispanic families, but also one of the highest rates of families living in poverty in the nation (Brudereck, 2012). Evaluators must possess a thorough knowledge of Hispanic culture, impoverished families, and the unique interactions between the two that are specific to the Reading area. School psychologists in the BCIU EI program at the time the study occurred each had, at minimum, six years of experience working in the Reading area.

Additionally, a minority of students receive a diagnosis of autism from another agency prior to their initial evaluation by the BCIU EI program. Outside diagnoses originate from sources including behavioral health rehabilitation service providers, developmental pediatricians, psychiatrists, and licensed psychologists.

These diagnoses include autistic disorder, pervasive developmental disorder, not otherwise specified, and Asperger's disorder, as defined in the *DSM-IV-TR* (APA, 2000). No students were classified under the new diagnostic category of autism spectrum disorder, as defined in the *DSM-5* (APA, 2013), as students were diagnosed prior to the manual's publication. Table 3 describes the number of students in the current sample who received an educational classification of autism through the BCIU EI program (n = 88) and the number of students who received a diagnosis from other service providers (n = 21).

Table 3

Source of Autism Diagnoses

Source	n	%
BCIU EI Program	88	80.8
ADOS	(85)	(78.0)
CARS-2	(3)	(2.8)
Other Service Providers	21	19.3
Total	109	100

Procedure

The primary research method employed was record reviews examining archival data from the BCIU EI program. Archival student data are stored in the EI program's IEP Writer database and in students' files.

Permission to duplicate and use the M-CHAT for the purposes of this study was obtained from the primary author of the M-CHAT, Dr. Diana Robins (see Appendix C). Permission to use students' archival data from the BCIU EI program was obtained from the director of the Office of Early Childhood and Student Services (currently assistant executive director of the BCIU), Dr. Jill Hackman (see Appendix D). The proposed research study was approved by the Indiana University of Pennsylvania Institutional Review Board for the Protection of Human Subjects.

M-CHAT forms and demographic information were gathered by BCIU EI employee volunteers. Each student record received a numerical code (Participant 1, Participant 2, etc.) and identifying information was removed. Results were then analyzed by the primary researcher. These archival data were anonymous in nature when examined.

Data Analysis

The following section will review the research questions, hypotheses, and the statistical analyses used to examine each. The IBM SPSS Statistics program, Version 20, was used to complete data analysis.

Research Question 1

How does the M-CHAT perform as an autism screening measure in the current sample of preschool-age students

(30-72 months) based on the traditional scoring procedures established for 16-30 month olds? It was hypothesized that there would be a high rate of false positives (or low PPV; number of students who failed the M-CHAT but do not have autism), as indicated by previous research (Robins, 2008; Robins, Fein, & Barton, 1999, Robins, Fein, Barton, & Green, 2001). It was also hypothesized that the sensitivity of the M-CHAT would be lower than reported in the initial validation study, as research has suggested that the M-CHAT is less sensitive in older children (Eaves, Wingert, & Ho, 2006; Snow & Lecavalier, 2008).

Descriptive statistics were generated to examine the numbers of students with autism and students with other developmental delays who failed and passed the M-CHAT based on traditional scoring procedures. The numbers of true positives, false positives, true negatives, and false negatives in the current sample were determined. Sensitivity, specificity, PPV, and negative predictive value (NPV) of the M-CHAT for the current sample and scoring method were also calculated. Receiver operator characteristic (ROC) analysis was conducted to depict the sensitivity and specificity of the M-CHAT in the form of a curve. As the assumptions of ROC analysis are the use of a dichotomous outcome measure (autism vs. other developmental delays) and a continuous predictor variable (number of items failed on the M-CHAT), separate ROC curves were constructed for the two scoring criteria of the M-CHAT. The first ROC curve included the six critical items (with a cutoff score of two failed items) and the second ROC curve included all items from the M-CHAT (with a cutoff score of three failed items). The numerical interpretation of the ROC curve, the area under the curve (AUC), was determined for each scoring criterion, with values closest to 1 indicating better predictive validity for the screening measure.

Research Question 2

How does the M-CHAT perform as an autism screening measure in the current sample of preschool-age students (30-72 months) based on the Best7 scoring procedures suggested for 16-30 month olds? Although there is no published literature regarding the Best7 scoring algorithm at this time, it was hypothesized that there would be a more moderate rate of false positives (number of students who failed the M-CHAT but do not have autism), as the Best7 criteria were suggested in order to decrease the high rate of false positives previously found with traditional scoring. It was also hypothesized that the sensitivity of the M-CHAT would be lower than reported in the initial validation of the study, due to the age range of the current sample.

Descriptive statistics were generated to examine the numbers of students with autism and students with other developmental delays who failed and passed the M-CHAT based on Best7 scoring procedures. The numbers of true positives, false positives, true negatives, and false negatives in the current sample were determined. Sensitivity, specificity, PPV, and NPV of the M-CHAT for the current sample and scoring method were also calculated. ROC analysis was conducted to depict the sensitivity and specificity of the M-CHAT in the form of a curve. As discussed in Research Question 1, separate ROC curves were constructed for the two scoring criteria of the M-CHAT. The first ROC curve included the seven critical items (with a cutoff score of two failed items) and the second ROC curve included all items from the M-CHAT (with a cutoff score of three failed items). The AUC was again determined for each scoring criterion.

Research Question 3

What are the differences in the M-CHAT's performance as an autism screening measure in the current sample, based on the use of the traditional scoring procedures vs. the Best7 scoring procedures? It was hypothesized that fewer false positives would be identified with the use of the Best7 scoring procedures, while the sensitivity would remain relatively constant, as the intention of the Best7 scoring algorithm was to increase PPV without significantly affecting sensitivity and specificity (Robins et al., 2010).

The M-CHAT's specificity, sensitivity, PPV, and NPV generated from Research Questions 1 and 2 were compared descriptively to examine the differences between each. To determine statistically significant differences in the AUCs between traditional and Best7 scoring methods, a z-test was conducted following recommendations delineated by Hanley and McNeil (1983).

Research Question 4

Which items on the M-CHAT are most predictive of preschool-age students (30-72 months) in the current sample receiving an educational classification of autism? Critical items were originally identified as children's ability to take an interest in other children, use their index finger to point to objects of interest, bring objects to their parents to show them something, imitate their parents, respond to their name, and follow a parent's point in order to look at something across the room (Robins et al., 2001). When the Best7 scoring algorithm was suggested the critical item of imitation was eliminated and two critical items were added, including engaging in pretend play and wondering if the child was deaf (Robins et al., 2010). It was hypothesized that similar items would be identified as strong predictors of autism; however, given the age of the sample it was also hypothesized that more items than these six to seven would be identified as strongly predictive of autism.

Stepwise discriminant function analysis was conducted to determine which items on the M-CHAT were most predictive of an educational classification of autism in the current sample. Assumptions that must be met for discriminant function analysis include interval or dichotomous variables, mutually exclusive and pre-defined categories, non-multicollinearity, and homoscedasticity (or homogeneity of variance).

Research Question 5

Does a principal components analysis of parent ratings indicate that the M-CHAT is measuring two components of autism (i.e., social communication deficits and unusual/repetitive behaviors), as per the newly recommended autism identification model, or will a different solution best represent findings from the current sample? It was hypothesized that the M-CHAT does not measure separate factors for social and communication deficits, based on factor analysis, cluster analysis, and principal components analysis conducted on other autism evaluation instruments (Constantino et al., 2004; Goldstein & Naglieri, 2010; Gotham, Risi, Pickles, & Lord, 2007; Robertson, Tanguay, L'Ecuyer, Sims, & Waltrip, 1999).

Principal components analysis was used to determine whether a two-factor model best fits the M-CHAT, or whether a different model more accurately defined the construct of this screening measure based on the current sample's performance. Assumptions that must be met for this analysis include the use of interval, ordinal, or dichotomous data and the presence of variables demonstrating a correlational relationship with each other. Table 4 provides an overview of the research questions and hypotheses, along with the variables analyzed, the statistical procedures used, and the assumptions that must be met for each.

Summary

The purpose of this study was to examine the utility of the M-CHAT in a preschool-age special education sample. This study sought to determine which items on the M-CHAT were most associated with students' receiving an educational classification of autism, as discriminated from

Table 4

Research Questions, Hypotheses, Variables, Statistical Analyses, and Assumptions

Research Questions	Hypotheses	Variables	Statistical Analyses	Statistical Assumptions
1. How does the M-CHAT perform as an autism screening measure in the current sample of preschool-age students (30-72 months) based on the traditional scoring procedures established for 16-30 month olds?	There will be a high rate of false positives (or low PPV; number of students who failed the M-CHAT but do not have autism). Sensitivity of the M-CHAT will be lower than reported in the initial validation study.	Educational classification (ASD/no ASD), M-CHAT outcome (pass/fail) based on traditional scoring algorithm	Descriptive statistics Frequency distributions ROC analysis	1. Continuous predictor variable 2. Dichotomous outcome variable
2. How does the M-CHAT perform as an autism screening measure in the current sample of preschool-age students (30-72 months) based on the Best7 scoring procedures suggested for 16-30 month olds?	There will be a moderately high rate of false positives (or low PPV; number of students who failed the M-CHAT but do not have autism). Sensitivity of the M-CHAT will be lower than reported in the initial validation study.	Educational classification (ASD/no ASD), M-CHAT outcome (pass/fail) based on Best7 scoring algorithm	Descriptive statistics Frequency distributions ROC analysis	1. Continuous predictor variable 2. Dichotomous outcome variable
3. What are the differences in the M-CHAT's performance as an autism screening measure in the current sample, based on the use of the traditional scoring procedures vs. the Best7 scoring procedures?	Fewer false positives will be identified with the use of the Best7 scoring procedures, while sensitivity will remain relatively constant.	Educational classification (ASD/no ASD), M-CHAT outcome (pass/fail), Scoring method (traditional/ Best7)	Examination of descriptive statistics, z-test examining significant differences between AUCs	1. AUCs from previously calculated ROC analyses 2. Continuous predictor variable 3. Dichotomous outcome variable

Table 4 (Continued)

Research Questions, Hypotheses, Variables, Statistical Analyses, and Assumptions

Research Questions	Hypotheses	Variables	Statistical Analyses	Statistical Assumptions
4. Which items on the M-CHAT are most predictive of preschool-age students (30-72 months) in the current sample receiving an educational classification of autism?	Similar items will be identified as strong predictors of autism as were identified by the M-CHAT validation study; however, it is expected that more items than the original 6-7 will be identified as strongly predictive of autism.	Educational classification (ASD/no ASD), item responses for each of the 23 items on the M-CHAT (pass/fail)	Stepwise discriminant function analysis	1. Interval or dichotomous variables 2. Categories are mutually exclusive and exhaustive 3. Categories are defined a priori 3. Multicollinearity 4. Homoscedasticity
5. Does a principal components analysis of parent ratings indicate that the M-CHAT is measuring two components of autism (i.e., social communication deficits and unusual/repetitive behaviors), as per the newly recommended autism identification model, or will a different solution best represent findings from the current sample?	There will not be sufficient evidence to support separate factors for social and communication deficits on the M-CHAT.	Item responses for each of the 23 items on the M-CHAT (pass/fail)	Principal components analysis	1. Interval, ordinal, or dichotomous data 2. Variables demonstrate a relationship with each other

students with other developmental delays. In addition, this study sought to determine whether a two-factor model of autism identification (i.e., social communication deficits and unusual/repetitive behaviors) or a different model best describes student performance on the M-CHAT.

To complete this research, archival data were gathered for 266 participants (n = 133 students with autism and n = 133 students with other developmental delays), including M-CHAT results and demographic information (age, sex, race, and educational classification). Complete data sets were available for 222 of the original 266 students. Participants' identifying information was removed from records and results were analyzed anonymously. Descriptive statistics, ROC analysis, discriminant function analysis, and principal components analysis were used to examine the utility of the M-CHAT in a preschool-age special education sample.

CHAPTER IV

DATA AND ANALYSES

In this chapter the research questions for the current study will be reviewed and data analysis procedures used for each will be discussed. The results for each question will be presented along with a discussion of whether the hypotheses for each research question were accepted or rejected based on the available data.

Research Question 1 Results

The first research question was: How does the Modified Checklist for Autism in Toddlers (M-CHAT; Robins, Fein, & Barton, 1999) perform as an autism screening measure in the current sample of preschool-age students (30-72 months) based on the traditional scoring procedures established for 16-30 month olds? It was hypothesized that there would be a high rate of false positives (or low Positive Predictive Value [PPV]; number of students who failed the M-CHAT but do not have autism), as indicated by previous research (Robins, 2008; Robins, Fein, Barton, & Green, 2001). It was also hypothesized that the sensitivity of the M-CHAT would be lower than reported in the initial validation study, as research has suggested that the M-CHAT is less sensitive in older children (Eaves, Wingert, & Ho, 2006; Snow & Lecavalier, 2008).

Descriptive statistics were generated to examine the number of students with autism and students with other developmental delays who failed and passed the M-CHAT based on traditional scoring procedures. The numbers of true positives, false positives, true negatives, and false negatives in the current sample were determined. As indicated in Table 5, the traditional scoring method (failure of the M-CHAT = failure of any three items or any two critical items) resulted in correct classification of 138 out of 222 students, or 62%. Correct classifications included students who failed the M-CHAT and received an educational classification of autism and students who passed the M-CHAT and did not receive an educational classification of autism. Thirty-eight percent of students (n = 84) were misclassified. Misclassifications included students who failed the M-CHAT and did not receive an educational classification of autism and students who passed the M-CHAT and did receive an educational classification of autism. Of students with autism, 70 failed the M-CHAT, while 39 passed. Thus, the number of true positives was 70 (64.2%) and the number of false negatives was 39 (35.7%). Of students who did not have autism, 45 failed the M-CHAT, while 68 passed. Thus, the number of false positives was 45 (39.8%) and the number of true negatives was 68 (60.2%).

Table 5

Pass/Fail Rates on the M-CHAT Using Traditional Scoring

	Diagnosis	
	ASD	Not ASD
Failed M-CHAT	**70** (64.2%)	*45* (39.8%)
Passed M-CHAT	*39* (35.7%)	**68** (60.2%)
Total	109	113

Note. Boldface = correct classifications; Italics = incorrect classifications; ASD = autism spectrum disorder; M-CHAT = Modified Checklist for Autism in Toddlers.

The sensitivity, specificity, PPV, and negative predictive value (NPV) of the M-CHAT for the current sample and scoring method were also calculated. The sensitivity (number of true positive cases [70] divided by the total number of students with autism [109]) was .64 for the current sample. The specificity (number of true negative cases [68] divided by total number of students without autism [113]) was .60. The PPV for the M-CHAT (number of true positive cases [70] divided by the total number of students who failed the M-CHAT [115]) was .61. Finally, the NPV (number of true negative cases [68] divided by total number of students who passed the M-CHAT [107]) was .64.

Receiver operator characteristic (ROC) analysis was conducted to depict the sensitivity and specificity of the M-CHAT in the form of a curve. As the assumptions of ROC analysis are the use of a dichotomous outcome measure (autism vs. other developmental delays) and a continuous predictor variable (number of items failed on the M-CHAT), separate ROC curves were constructed for the two scoring criteria of the M-CHAT. The first ROC curve included the six critical items (with a cutoff score of two failed items) and the second ROC curve included all items from the M-CHAT (with a cutoff score of any three failed items). The numerical interpretation of the ROC curve, the area under the curve (AUC), was determined for each scoring criterion, with values closest to 1 indicating better predictive validity for the screening measure, and numbers closer to 0.5 indicating random chance.

As indicated in Table 6, the AUC was .697 when the six critical items were included in the ROC analysis. The null hypothesis of ROC analysis is that the true AUC is equal to 0.5. As the difference between the expected value of 0.5 and the true value of the AUC was statistically significant ($p < .001$), the null hypothesis can be rejected. It can be stated that the relationship between ratings on the M-CHAT and participants' educational classification did not happen by chance. The AUC was .679 when all 23 items were included in the ROC analysis ($p < .001$), indicating that the relationship between ratings and educational classification was statistically significant and not due to chance.

Table 6

ROC Results for Critical Items and All Items on the M-CHAT

	AUC	*p*	*SE*	95% CI
Critical Items	.697	.000*	.036	.628 - .767
All Items	.679	.000*	.036	.608 - .749

Note. AUC = area under the curve; SE = standard error of measurement; CI = confidence interval.
*$p < .001$

Figures 1 and 2 depict the ROC curves generated for the six critical items and all items on the M-CHAT, respectively. The diagonal line indicates random chance, while the area between the two lines (AUC) depicts all possible combinations of specificity and sensitivity for the M-CHAT.

In summation, it was hypothesized that there would be a high rate of false positives (or low PPV; number of students who failed the M-CHAT but do not have autism). It was also hypothesized that the sensitivity of the M-CHAT would be lower than reported in the initial validation study (Robins et al., 2001). Results of the current study

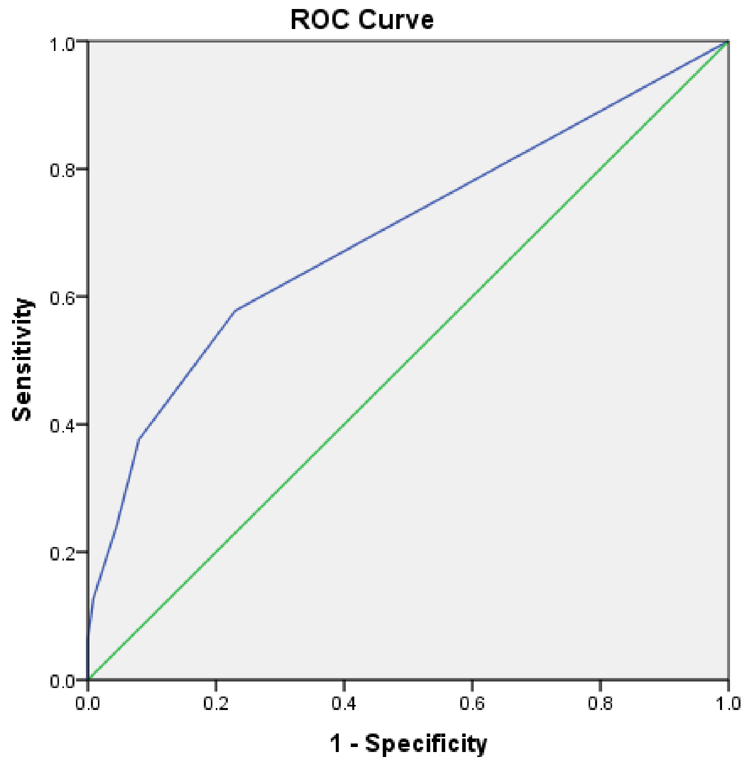

Figure 1. ROC curve including the six critical items on the M-CHAT.

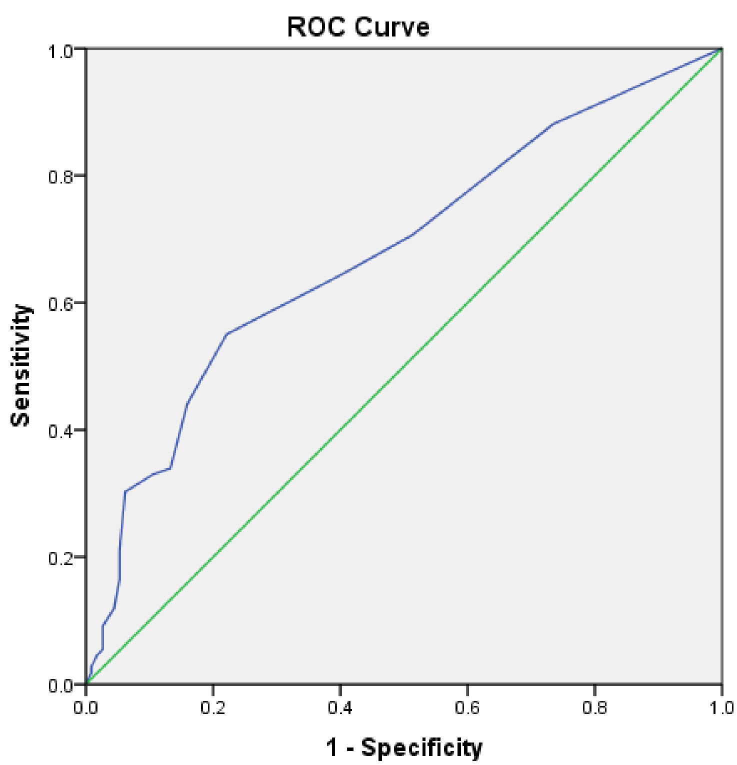

Figure 2. ROC curve including all 23 items on the M-CHAT.

indicate that the hypothesis was confirmed. A large number of false positives (students who failed the M-CHAT but do not have autism) were identified. Overall, 45 of 113 students who did not have autism failed the M-CHAT, or 39.8% of all students who did not have autism. Additionally, the sensitivity of the M-CHAT in the current study (.64) was much lower than the range of .87 - .97 that was reported in the initial validation study (Robins et al., 2001).

Research Question 2 Results

The second research question was: How does the M-CHAT perform as an autism screening measure in the current sample of preschool-age students (30-72 months) based on the Best7 scoring procedures suggested for 16-30 month olds? It was hypothesized that there would be a more moderate rate of false positives (number of students who failed the M-CHAT but do not have autism), as the Best7 criteria were developed in order to decrease the high rate of false positives previously reported with traditional scoring. It was also hypothesized that sensitivity of the M-CHAT would again be lower than reported in the initial validation of the study, due to the age range of the current sample.

Descriptive statistics were generated to examine the numbers of students with autism and students with other developmental delays who failed and passed the M-CHAT based on Best7 scoring procedures. The numbers of true positives, false positives, true negatives, and false negatives in the current sample were determined. As indicated in Table 7, using the Best7 scoring method (failure of the M-CHAT = failure of any three items or any two Best7 critical items) the M-CHAT correctly classified 138 out of 222 students, or 62%. Thirty-eight percent (n = 84) of students were misclassified. Of students with autism, 70 failed the M-CHAT, while 39 passed. Thus, the number of true positives was 70 (64.2%) and the number of false negatives was 39 (35.7%). Of students who did not have autism, 45 failed the M-CHAT, while 68 passed. Thus, the number of false positives was 45 (39.8%) and the number of true negatives was 68 (60.2%).

Sensitivity, specificity, PPV, and NPV of the M-CHAT for the current sample and scoring method were also calculated. The sensitivity was .64 for the current sample while the specificity was .60. The PPV for the M-CHAT was .61 and the NPV was .64.

Table 7

Pass/Fail Rates on the M-CHAT Using Best7 Scoring

	Diagnosis	
	ASD	Not ASD
Failed M-CHAT	**70** (64.2%)	*45* (39.8%)
Passed M-CHAT	*39* (35.7%)	**68** (60.2%)
Total	109	113

Note. Boldface = correct classifications; Italics = incorrect classifications; ASD = autism spectrum disorder; M-CHAT = Modified Checklist for Autism in Toddlers.

ROC analysis was conducted to depict the sensitivity and specificity of the M-CHAT in the form of a curve. Similar to the procedure used in Research Question 1, separate ROC curves were constructed for the two scoring criteria of the M-CHAT. The first ROC curve included the Best7 critical items and the second ROC curve included all items from the M-CHAT. The AUC was again determined for each scoring criterion.

As demonstrated in Table 8, the AUC was .712 when the Best7 critical items were included in the ROC analysis. The null hypothesis of ROC analysis is that the true AUC is equal to 0.5. As the difference between the expected value of 0.5 and the true value of the AUC was statistically significant ($p < .001$), the null hypothesis can be rejected. It can be stated that the relationship between ratings on the M-CHAT and participants' educational classification did not happen by chance. The AUC was .679 when all 23 items were included in the ROC analysis ($p < .001$), again indicating that the relationship between ratings and educational classification was statistically significant and not due to chance.

Table 8

ROC Results for Best7 Items and All Items on the M-CHAT

	AUC	*p*	*SE*	95% CI
Best7 Items	.712	.000*	.035	.644 - .781
All Items	.679	.000*	.036	.608 - .749

Note. AUC = area under the curve; SE = standard error of measurement; CI = confidence interval.
*$p < .001$

Figures 3 and 4 depict the ROC curves generated for the Best7 items and all items on the M-CHAT, respectively. The diagonal line indicates random chance, while the area between the two lines (AUC) depicts all possible combinations of specificity and sensitivity for the M-CHAT.

It was hypothesized that there would be a more moderate rate of false positives (number of students who failed the M-CHAT but do not have autism), as the Best7 criteria were developed in order to decrease the high rate of false positives previously reported with traditional scoring. It was also hypothesized that sensitivity of the

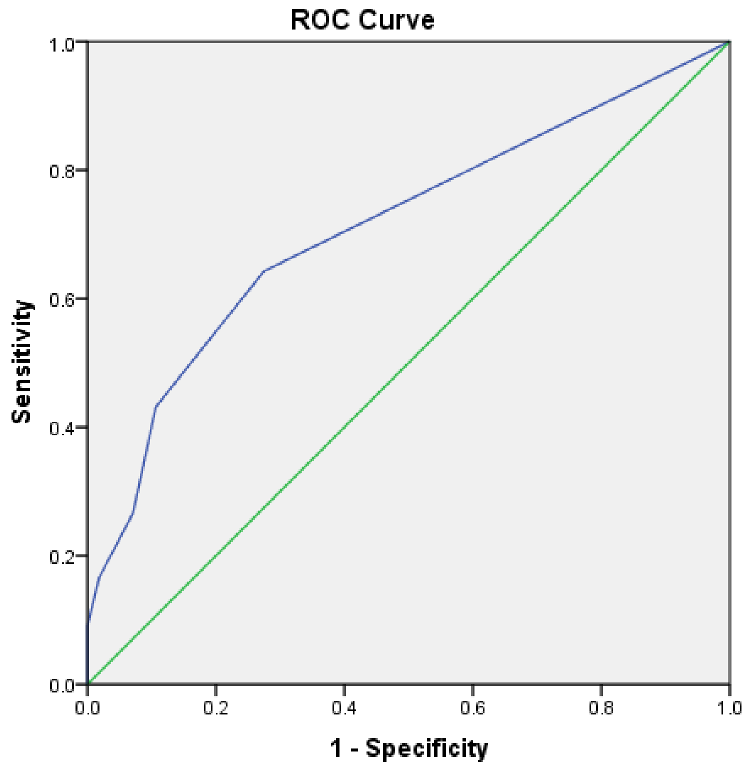

Figure 3. ROC curve including the Best7 items on the M-CHAT.

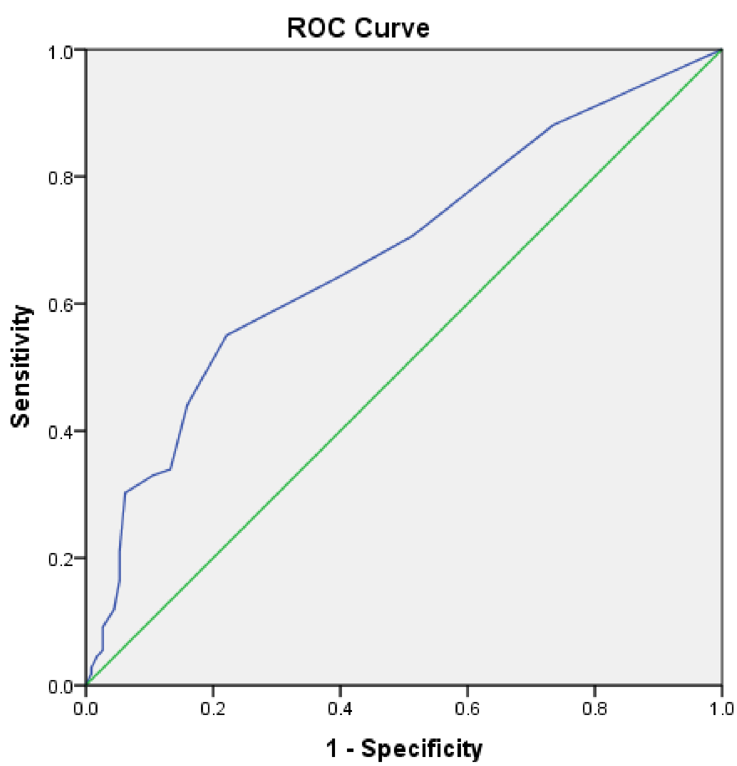

Figure 4. ROC curve including all 23 items on the M-CHAT.

M-CHAT would again be lower than reported in the initial validation of the study, due to the age range of the current sample. Results of the current study indicate that the hypothesis of generating fewer false positives with the Best7 scoring method was rejected. A similarly large number of false positives were identified with the Best7 scoring method as were identified with the traditional scoring method. Overall, 45 of 113, or 39.8% of students who did not have autism failed the M-CHAT using the Best7 scoring method. The hypothesis that the sensitivity of the M-CHAT would again be lower than in the initial validation study, however, was confirmed. The sensitivity of the M-CHAT in the current study (.64) was much lower than the range of .87 - .97 reported in the initial validation study (Robins et al., 2001).

Research Question 3 Results

The third research question was: What are the differences in the M-CHAT's performance as an autism screening measure in the current sample, based on the use of the traditional scoring procedures vs. the Best7 scoring procedures? It was hypothesized that fewer false positives would be identified with the use of the Best7 scoring procedures, while the sensitivity would remain relatively constant; as the intention of the Best7 scoring algorithm was to increase PPV without significantly affecting sensitivity and specificity (Robins et al., 2010).

Pass/fail rates along with the M-CHAT's specificity, sensitivity, PPV, and NPV generated from Research Questions 1 and 2 were compared descriptively to examine the differences between each. When the M-CHAT was rescored with the Best7 method there were no differences in the number of students who had passed and failed the M-CHAT using the traditional scoring method. With both scoring methods the M-CHAT correctly classified 138 out of 222 students, or 62%. As demonstrated by Table 9, every student who passed the M-CHAT using the traditional scoring method also passed using the Best7 method; likewise all students who failed the M-CHAT using the traditional scoring method failed using the Best7 method. There were no differences between the rates of true positives (n = 70), false positives (n = 45), true negatives (n = 68), or false negatives (n = 39). Of students with autism, 70 failed the M-CHAT, while 39 passed. Of students who did not have autism, 45 failed the M-CHAT, while 68 passed (see Table 11). As indicated by Table 12, there were no differences in the numbers of students who passed and failed the M-CHAT, the sensitivity (.64), specificity

(.60), PPV (.61), and NPV (.64) were also the same for both scoring methods.

Table 9

Comparison of Pass/Fail Rates on the M-CHAT for Traditional vs. Best7 Scoring Procedures

Scoring	Traditional M-CHAT		Best7 M-CHAT	
	ASD	Not ASD	ASD	Not ASD
Failed M-CHAT	**70**	*45*	**70**	*45*
Passed M-CHAT	*39*	**68**	*39*	**68**
Total	109	113	109	113

Note. Bold type = correct classifications; Italic type = incorrect classifications; ASD = autism spectrum disorder; M-CHAT = Modified Checklist for Autism in Toddlers.

Table 10

Comparison of Sensitivity, Specificity, Positive Predictive Values, and Negative Predictive Values for Traditional vs. Best7 Scoring Procedures

	Sensitivity	Specificity	PPV	NPV
Traditional	.64	.60	.61	.64
Best7	.64	.60	.61	.64

Note. PPV = positive predictive value; NPV = negative predictive value.

Although there were no differences in the actual rates of students who passed and failed the M-CHAT when different scoring methods were used, it was noted the AUCs from the previously conducted ROC analyses were slightly different

(.697 for traditional scoring and .712 for Best7 scoring). To determine statistically significant differences between the AUCs, a z-test was conducted following recommendations delineated by Hanley and McNeil (1983). With a difference of .015 between the two AUCs, results indicated a z-score of 0.3042. As indicated by Table 11, the difference between the AUCs was not significant as the z-score of 0.3042 was less than the critical value of 1.96 at an alpha level of .05.

Table 11

Comparison of the AUCs for Traditional vs. Best7 Scoring Procedures

AUC		Difference in AUC	*z*-score	Critical value	*p* (two-tailed)
Traditional	.697	.015	0.3042	1.96	0.760975
Best7	.712				

Note. AUC = area under the curve.
p < .05

In review, it was hypothesized that fewer false positives would be identified with the use of the Best7 scoring procedures, while the sensitivity would remain relatively constant; as the intention of the Best7 scoring algorithm was to increase PPV without significantly affecting sensitivity and specificity (Robins et al., 2010). This hypothesis was rejected, as no differences were found in the rate of false positives or in the sensitivity of the M-CHAT when the two scoring methods were compared.

Research Question 4 Results

The fourth research question was: Which items on the M-CHAT are most predictive of preschool-age students (30-72 months) in the current sample receiving an educational classification of autism? It was hypothesized that similar items would be identified as strong predictors of autism as the critical items identified in the traditional and Best7 scoring methods; however, given the age of the sample it was also hypothesized that more items than these six to seven would be identified as strongly predictive of autism.

Stepwise discriminant function analysis was conducted to determine which items on the M-CHAT were most predictive of students receiving an educational classification of autism in the current sample. Assumptions that must be met for discriminant function analysis include interval or dichotomous variables, mutually exclusive and pre-defined categories, non-multicollinearity, and homoscedasticity.

To meet the assumption of interval or dichotomous data for the predictor variables (parent responses to individual questions on the M-CHAT), students' scores of pass and fail were dummy coded in a dichotomous scheme where 0 = pass and 1 = fail. The categories for the analysis were mutually exclusive and had been defined a priori (autism vs. not autism).

The general rule of thumb for determining the presence of multicollinearity (meaning that the predictor variables should not be strongly correlated with each other) is that Pearson's correlation coefficients between the variables should be less than 0.8, and the variance inflation factor (VIF) for each variable should be less than or equal to 3.3 (Field, 2009). As demonstrated by Tables 12 and 13, all VIF statistics for the predictor variables were less than 3.3 and an examination of the point biserial correlation matrix (a special case of Pearson's correlations, used for dichotomous data) for the M-CHAT questions did not reveal any correlations greater than 0.8.

Homoscedasticity, or homogeneity of variance, refers to equal variance of the ranges for each of the predictor variables. As a dichotomous coding system was used for all predictor variables, all variables were coded as 0 or 1. Therefore, the amount of variance for each variable was equal, as each was confined to a range of 0 to 1 and could not vary beyond this pre-established range. Additionally, box plots were constructed for each question on the M-CHAT to ensure that there were no items that all students passed or all students failed, which would indicate zero variance.

Table 12

VIF and Tolerance Statistics for M-CHAT Questions

	Collinearity Statistics	
	Tolerance	VIF
Q1	.669	1.494
Q2	.522	1.916
Q3	.807	1.239
Q4	.647	1.545
Q5	.494	2.026
Q6	.439	2.280
Q7	.418	2.391
Q8	.615	1.627
Q9	.573	1.745
Q10	.742	1.347
Q11	.822	1.217
Q12	.641	1.561
Q13	.568	1.761
Q14	.546	1.832
Q15	.339	2.946
Q16	.580	1.725
Q17	.495	2.021
Q18	.648	1.544
Q19	.547	1.829
Q20	.687	1.456
Q21	.592	1.689
Q22	.541	1.848
Q23	.653	1.531

Note. VIF = variance inflation factor.

As demonstrated by Figure 5, it was noted that the vast majority of students passed item 16 (Does your child walk?), limiting the variance of this question. Item 16 was meant to be a buffer question, or a question unrelated to autism that most parents could answer positively regardless of how severely affected by autism their child was (Robins et al., 2001). As this item was unrelated to autism, it was not expected to significantly impact

Table 13

Point Biserial Correlation Matrix for M-CHAT Questions

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
Q1	1.00										
Q2	.22**	1.00									
Q3	.17*	.08	1.00								
Q4	.26**	.34**	.15*	1.00							
Q5	.01	.38**	.18*	.37**	1.00						
Q6	.19**	.29**	.13	.24**	.33**	1.00					
Q7	.24**	.42**	.07	.33**	.28**	.68**	1.00				
Q8	.13	.31**	.15*	.24**	.30**	.24**	.20**	1.00			
Q9	.12	.34**	.10	.28**	.44**	.33**	.29**	.25**	1.00		
Q10	.10	.26**	.07	.12	.13*	.19**	.09	.24**	.12	1.00	
Q11	.08	.16*	.03	.10	.04	.12	.14*	.04	.13	.08	1.00
Q12	.11	.12	.06	.25**	.16*	.15*	.15*	.19**	.30**	.10	.07
Q13	-.03	.23**	.14*	.26**	.50**	.25*	.28**	.26**	.47**	.16*	.02
Q14	.13	.20*	.06	.21**	.07	.17*	.18**	.14*	.22**	.35**	.04
Q15	.13	.39**	.10	.33**	.30**	.50*	.41**	.29**	.45**	.36**	.13
Q16	-.03	.03	.15*	.10	.05	.15*	.26**	.07	.12	-.05	-.10
Q17	.13	.28**	.08	.18**	.23**	.27**	.23**	.27**	.30**	.25**	.06
Q18	.01	.19**	.09	.05	.16*	.12	.14*	.29**	.14*	.18**	.18**
Q19	.19**	.38**	.04	.27**	.43**	.28**	.32**	.33**	.43**	.25**	.06
Q20	.10	.22**	.07	.07	.17*	.19**	.12	.03	.26**	.12	.14*
Q21	.19**	.30**	.17	.20**	.15*	.19**	.13	.24**	.17*	.29**	.00
Q22	.19**	.15*	.03	.11	.20**	.21**	.18**	.29**	.21**	.26**	.28**
Q23	.17*	.21**	.13	.29**	.09	.02	.16*	.01	.19**	.15*	-.05

*$p < .05$. **$p < .01$.

Table 13 (Continued)

Point Biserial Correlation Matrix for M-CHAT Questions

	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Q22	Q23
Q12	1.00											
Q13	.17*	1.00										
Q14	.09	.22**	1.00									
Q15	.31**	.38**	.46**	1.00								
Q16	.17**	.06	-.03	.04	1.00							
Q17	.30**	.24**	.43**	.62**	.17*	1.00						
Q18	.10	.23**	.02	.16*	.16*	.13	1.00					
Q19	.18**	.37**	.28**	.50**	.14*	.45**	.17*	1.00				
Q20	.03	.20**	.15*	.28**	-.05	.23**	.06	.25**	1.00			
Q21	.12	.14	.52**	.44**	.08	.36**	.16*	.25**	.19**	1.00		
Q22	.03	.22**	.19**	.25**	-.01	.26**	.43**	.21**	.32**	.24**	1.00	
Q23	.33**	.14*	.29**	.24**	.13	.31**	.12	.30**	.11	.24**	.09	1.00

*p < .05. **p < .01.

results, and was eliminated from the analysis. Thus, all assumptions for the use of discriminant function analysis were met.

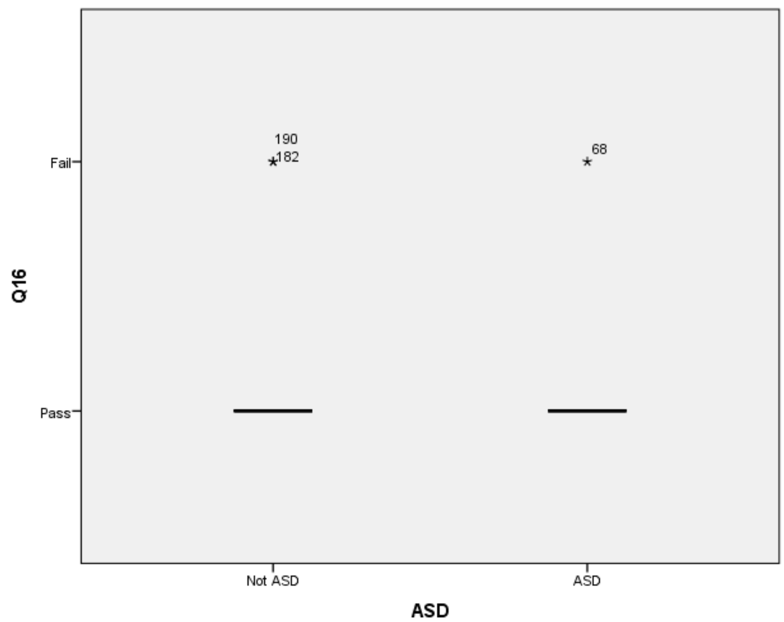

Figure 5. Box plot depicting pass/fail rates for question 16 on the M-CHAT for students with and without ASD (autism spectrum disorder), indicating that only three students failed this question.

As indicated in Table 14, discriminant function analysis using a stepwise procedure indicated that two variables best minimized Wilks's Lambda, or Λ: Question 13 ("Does your child imitate you? [e.g., you make a face - will your child imitate it?]"; Robins, Fein, & Barton, 1999), and Question 2 ("Does your child take an interest in other children?"). Based on the analysis, failure of Question 13 was the single best predictor of a child receiving an educational classification of autism, and failure of Question 2 was the next best predictor. With the inclusion of these two variables, Λ was reduced to .860. The model was statistically significant ($p < .001$), indicating that a two-step model was a good fit for the available data.

Table 14

Wilks's Lambda for Stepwise Discriminant Function Analysis

Step	Variable	Λ	*F*	*p*
1	Q13	.894	22.144	.000*
2	Q2	.860	15.061	.000*

*$p < .001$

The purpose of this research question was to determine whether new or different critical items could be identified for the current sample that would lead to an improved scoring method. As the two items identified as being strongly predictive of autism were already identified using the traditional scoring critical item procedures, no new critical items were identified. A further examination of results was then undertaken to determine whether failure of just one of these items would make a difference in scoring. It was found that of the 109 students with autism, 35 failed question 2 and 28 failed question 13. Of the 113 students without autism, 12 failed question 2 and five failed question 13. If students failed the M-CHAT based on the criterion of failing either one of these two items, an additional two students with autism would fail the screener; however, an additional three students without autism would have failed as well. Thus, using this criterion would potentially worsen the predictive validity of the M-CHAT in the current sample.

In addition, a review of parent responses indicated that no participant passed or failed the M-CHAT based on the use of critical items, neither for the traditional scoring method nor for the Best7 scoring method. All students passed or failed based on the cutoff criterion of failing three or more items. Analysis was then undertaken to examine the M-CHAT's performance using different cutoff criteria to determine whether predictive validity could be improved in this manner. As demonstrated by Table 15, attempts to improve the M-CHAT's performance as an autism screening measure by changing the cutoff criteria were not successful.

As shown below, if the cutoff criterion was changed to failing any two items, rather than three, an additional six students with autism could be correctly classified, raising the sensitivity to .70. However, this meant that an additional 13 students who did not have autism would also have failed the screening, which reduced the PPV to .57.

This created an unsatisfactory amount of false positives, which was already at an unacceptably high level. Changing the cutoff criteria to failing one item was also not favorable for similar reasons. While an additional 26 students with autism would be correctly classified, all but 30 of the 113 students without autism would have failed the screener. Raising the cutoff score to four items decreased the number of false positives, but lowered the sensitivity to .55.

Table 15

M-CHAT Performance with Varying Cutoff Criteria

Failure Cutoff	1 item	2 items	3 items[a]	4 items
True Positives	96	76	70	60
True Negatives	30	55	68	88
False Positives	83	58	45	25
False Negatives	13	33	39	49
Sensitivity	.88	.70	.64	.55
Specificity	.27	.49	.60	.78
PPV	.54	.57	.61	.71
NPV	.70	.63	.64	.64

Note. PPV = positive predictive value; NPV = negative predictive value.
a. Current cutoff recommendation.

Robins et al. (2001) originally identified critical items as children's ability to take an interest in other children, use their index finger to point to objects of interest, bring objects to their parents to show them something, imitate their parents, respond to their name, and follow a parent's point in order to look at something across the room. When the Best7 scoring algorithm was suggested, the critical item of imitation was eliminated and two critical items were added, including engaging in pretend play and wondering if the child was deaf (Robins et al., 2010). It was hypothesized that similar items would be identified as strong predictors of autism; however, given the age of the sample it was also hypothesized that more items than these six to seven would be identified as strongly predictive of autism.

The hypothesis was partially accepted, in that similar items were found to be highly predictive of autism; however, fewer items, rather than more, were found to be predictive of autism in the current model. Both items identified in the current analysis (ability to imitate and taking an interest in other children) were originally identified as being highly predictive of autism in the traditional scoring approach by outlined by Robins et al. (2001). On the other hand, only the item related to taking an interest in other children was identified as being predictive of autism using the Best7 scoring method. In addition, only two items were found to be significantly predictive of autism using the current sample, as opposed to six and seven items previously identified using the traditional and Best7 scoring approaches, respectively. Overall, the predictive validity of the M-CHAT could not be improved for the current sample by changing cutoff criteria or by identifying new critical items.

Research Question 5 Results

The fifth research question was: Does a principal components analysis of parent ratings indicate that the M-CHAT is measuring two components of autism (i.e., social communication deficits and unusual/repetitive behaviors), as per the newly recommended autism identification model, or will a different solution best represent findings from the current sample? It was hypothesized that the M-CHAT does not measure separate factors for social and communication deficits, based on factor analysis, cluster analysis, and principal components analysis conducted on other autism evaluation instruments (Constantino et al., 2004; Goldstein & Naglieri, 2010; Gotham, Risi, Pickles, & Lord, 2007; Robertson, Tanguay, L'Ecuyer, Sims, & Waltrip, 1999).

Assumptions that must be met for the use of principal components analysis include the use of interval, ordinal, or dichotomous data, and the presence of variables demonstrating a correlational relationship with each other.

As previously noted, students' scores of pass and fail on the M-CHAT were dummy coded in a dichotomous scheme where 0 = pass and 1 = fail to allow for the use of dichotomous data. As indicated in Tables 16 and 17, point biserial correlation matrices were constructed to determine the presence of one or more variables with correlations greater than 0.3 (Tabachnick & Fidell, 2007) for students with and without autism, respectively. An examination of the correlation matrices indicated multiple correlations greater than 0.3 between various items on the M-CHAT. It was noted that for students with autism, correlations between items 16 and 14 and items 16 and 21 could not be generated due to item 16 being constant. As discussed under Research Question 4, item 16 is not an autism related question and its removal was not expected to significantly impact results. Thus, item 16 was not included in the analyses for either group.

Principal components analyses were initially conducted separately for students with and without autism. As dissimilar items on the M-CHAT loaded onto the factors for each group (see discussion below), the groups were not combined for analysis.

Table 16

Point Biserial Correlation Matrix for M-CHAT Questions for Students with Autism

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
Q1	1.00										
Q2	.14	1.00									
Q3	.12	.03	1.00								
Q4	.33**	.35**	.04	1.00							
Q5	.01	.36**	.06	.30**	1.00						
Q6	.20	.29**	.18	.21*	.41**	1.00					
Q7	.18	.50**	.07	.37**	.28**	.60**	1.00				
Q8	.08	.25**	.13	.31**	.31**	.27**	.18	1.00			
Q9	.25**	.38**	.05	.26**	.41**	.32**	.30**	.22*	1.00		
Q10	.16	.21*	.11	.14	.08	.28**	.14	.14	.05	1.00	
Q11	.14	.23*	.15	.22*	.08	.10	.14	.06	.18	.21*	1.00
Q12	.10	.18	.05	.25**	.03	.15	.13	.11	.16	.09	.10
Q13	.01	.22*	.07	.25*	.44**	.27**	.28**	.22*	.43**	.16	.01
Q14	.19	.24*	.14	.18	-.03	.18	.17	.14	.18	.44**	.06
Q15	.10	.35**	.14	.32**	.20*	.50**	.36**	.27**	.38**	.38**	.18
Q16	-.03	.14	-.02	-.04	-.06	-.06	.16	-.04	-.04	-.05	-.09
Q17	.07	.21*	-.01	.11	.13	.18	.15	.14	.24*	.23*	.17
Q18	.00	.24*	.06	.11	.22*	.05	.17	.27**	.09	.20*	.12
Q19	.16	.35**	-.05	.27**	.45**	.35**	.31**	.26**	.39**	.29**	.09
Q20	.03	.13	-.02	.03	.16	.20*	.12	-.03	.27**	.16	.20*
Q21	-.04	.24*	.21*	.13	.09	.11	.05	.25*	.04	.33**	.01
Q22	.16	.16	.09	.16	.25**	.23*	.20*	.34**	.28**	.28**	.33**
Q23	.16	.20*	.04	.28**	-.03	-.08	.18	-.07	.14	.18	.01

**p* < .05. ***p* < .01.

Table 16 (Continued)

Point Biserial Correlation Matrix for M-CHAT Questions for Students with Autism

	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Q22	Q23
Q12	1.00											
Q13	.04	1.00										
Q14	.01	.17	1.00									
Q15	.28**	.35**	.39**	1.00								
Q16	-.02	-.06	a	-.07	1.00							
Q17	.27**	.17	.42**	.54**	.19	1.00						
Q18	.06	.25*	-.01	.15	.16	.08	1.00					
Q19	.02	.33**	.27**	.43**	.15	.39**	.25*	1.00				
Q20	-.04	.24*	.14	.28**	-.05	.23*	.05	.19	1.00			
Q21	.01	.09	.59**	.32**	a	.23*	.11	.17	-.02	1.00		
Q22	-.04	.30**	.19	.30**	-.08	.35**	.45**	.31**	.33**	.19	1.00	
Q23	.30**	.08	.35**	.19	.15	.27**	.11	.22*	-.03	.16	.14	1.00

a. Correlation could not be computed due to one of the variables being constant.
*$p < .05$. **$p < .01$.

Table 17

Point Biserial Correlation Matrix for M-CHAT Questions for Students without Autism

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
Q1	1.00										
Q2	.40**	1.00									
Q3	.20*	.17	1.00								
Q4	.24*	.20*	.33**	1.00							
Q5	.05	.27**	.43**	.48**	1.00						
Q6	.22*	.18	.07	.24**	.05	1.00					
Q7	.37**	.11	.08	.11	.07	.80**	1.00				
Q8	.19*	.37**	.18	.07	.27**	.16	.20*	1.00			
Q9	-.05	.12	.21*	.23*	.41**	.30**	.14	.30**	1.00		
Q10	.05	.31**	.02	.04	.21*	.03	-.03	.35**	.25**	1.00	
Q11	.03	-.01	-.09	-.16	-.11	.11	.10	-.01	-.00	-.08	1.00
Q12	.14	-.06	.21*	.23*	.41**	.12	.14	.30**	.66**	.09	-.00
Q13	-.07	-.07	.33**	.16	.48**	.06	.08	.34**	.50**	.16	-.07
Q14	.08	-.05	-.05	.16	.12	.06	.08	.07	.23*	.17	-.07
Q15	.25**	.32**	.08	.25**	.33**	.43**	.37**	.32**	.56**	.35**	-.01
Q16	-.04	-.05	.27*	.30**	.24**	.39**	.43**	.17	.40**	-.06	-.10
Q17	.25**	.32**	.22*	.25**	.33**	.34**	.27**	.45**	.36**	.27**	-.15
Q18	.02	.12	.12	-.04	.09	.20*	.12	.31**	.27**	.16	.25**
Q19	.28**	.32**	.16	.19*	.25**	.35**	.20*	.42**	.46**	.17	-.03
Q20	.21*	.27**	.19*	.07	.04	.08	.00	.08	.13	.04	.01
Q21	.41**	.35**	.13	.28**	.21*	.27**	.23*	.21*	.41**	.23*	-.03
Q22	.22*	.11	-.01	.02	.11	.16	.16	.22*	.11	.23*	.21*
Q23	.20*	.16	.33**	.26**	.19*	.10	.04	.09	.23*	.10	-.16

*$p < .05$. **$p < .01$.

Table 17 (Continued)

Point Biserial Correlation Matrix for M-CHAT Questions for Students without Autism

	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Q22	Q23
Q12	1.00											
Q13	.50**	1.00										
Q14	.23*	.19*	1.00									
Q15	.36**	.25**	.57**	1.00								
Q16	.40**	.30**	-.03	.21*	1.00							
Q17	.36**	.30**	.41**	.76**	.21*	1.00						
Q18	.15	.26**	.07	.19*	.16	.21*	1.00					
Q19	.46**	.33**	.19*	.53**	.16	.51**	.09	1.00				
Q20	.13	-.07	.07	.14	-.04	.12	.07	.26**	1.00			
Q21	.25**	.20*	.41**	.63**	.14	.54**	.22*	.33**	.47**	1.00		
Q22	.11	.11	.20*	.19*	.04	.12	.42**	.08	.32**	.30**	1.00	
Q23	.38**	.15	.14	.24*	.13	.32**	.12	.37**	.27**	.31**	.03	1.00

*p < .05. **p < .01.

Principal Components Analysis for Students with Autism

Prior to analysis, multiple criteria were examined in order to determine how many factors to retain for rotation. The Kaiser criterion suggests that all factors with eigenvalues greater than 1 should be retained; however, this rule often leads to an overestimation of the true number of factors (O'Connor, 2000). As shown in Table 18, a total of eight factors with eigenvalues greater than 1 were identified when this rule was applied. Cattell's scree test (Cattell, 1966) was also used as part of the decision-making process. This method employs visual examination of eigenvalues as plotted on a line graph. The decision regarding factor retention is based on the point at which the slope of the line bisecting the greatest number of components appears to significantly change. Because this method requires "eyeball searches of plots for sharp demarcations between the eigenvalues ... the reliability of scree plot interpretations is low, even among experts" (O'Connor, 2000, p. 396). As illustrated by Figure 6, examination of the scree plot indicated that two factors should be retained in the current analysis.

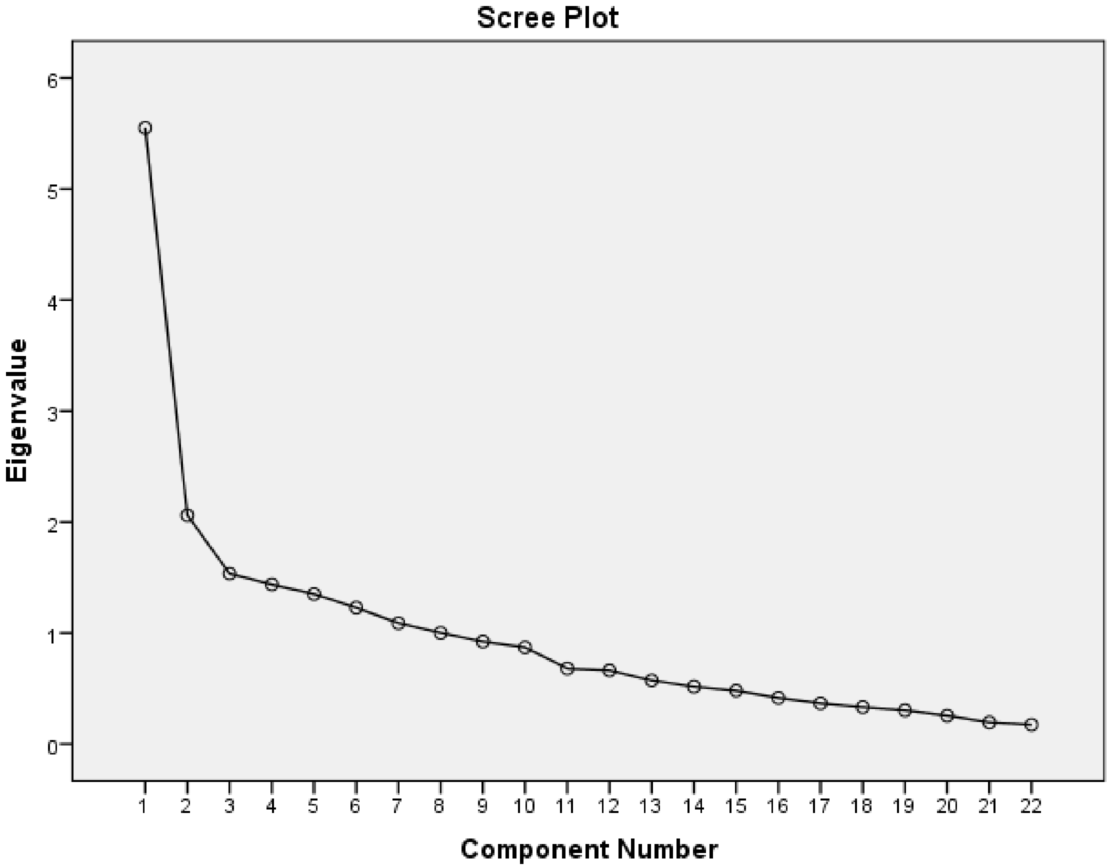

Figure 6. Scree plot of eigenvalues for students with autism indicating significant change of slope after two factors.

As the previous two decision-making procedures indicated dissimilar numbers of factors to retain (eight and two, respectively) and each method has notable flaws, a third procedure was utilized. Parallel analysis was used to compare the eigenvalues calculated from the actual data analysis with randomly generated eigenvalues (Horn, 1965). Per this method, factors are retained as long as the eigenvalues from the actual data set exceed the corresponding eigenvalues from the randomly generated data set (O'Connor, 2000). The software program Monte Carlo PCA for Parallel Analysis (Watkins, 2000) was used to conduct this analysis. As demonstrated by Table 18, a comparison of actual and randomly generated eigenvalues suggested that two factors should be retained in the current analysis.

Table 18

Actual and Randomly Generated Eigenvalues Used for Parallel Analysis for Students with Autism

Factor	Actual	Parallel Analysis
1	5.55	1.92
2	2.06	1.75
3	*1.54*	*1.62*
4	1.44	1.51
5	1.35	1.42
6	1.23	1.33
7	1.09	1.25
8	1.00	1.17
9	.92	1.10
10	.87	1.03
11	.68	.97
12	.66	.91
13	.57	.85
14	.52	.79
15	.48	.74
16	.42	.68
17	.37	.63
18	.33	.58
19	.30	.52
20	.26	.47
21	.20	.42
22	.17	.35

Note. Italics indicate the point at which actual eigenvalues are no longer greater than randomly generated eigenvalues. Random eigenvalues calculated using n = 109, number of variables = 22, replications = 1000.

Agreement between parallel analysis and scree plot examination suggested that a two-factor solution should be used for the current data set. Thus, principal components analysis limited to two factors with a varimax (orthogonal) rotation method was conducted for parent ratings on the M-CHAT for students with autism. As demonstrated in Table 19, results from this analysis indicated that a total of 34.6% of the variance was explained using a two-factor solution.

The first factor identified explained 20.39% of the total variance and included items related to social interaction, joint attention, play skills, nonverbal communication, and unusual behaviors; all deficits of and characteristics related to autism. This factor was named Core Features of Autism. Using the criterion of factor loadings greater than .40, this factor included items 2, 4, 5, 6, 7, 8, 9, 13, 15, 18, 19, 20, and 22 on the M-CHAT. These items addressed the following skills: Taking an interest in other children, enjoying playing games such as peek-a-boo or hide and seek, engaging in pretend play, using a distal point for the purpose of asking for something, using a distal point for the purpose of indicating interest, engaging in functional play, bringing objects to a parent for the purpose of showing, engaging in imitation, following another's point, making unusual finger movements near his/her face, attracting attention to

Table 19

Total Variance Explained for Principal Components Analysis Using a Two-Factor Solution for Students with Autism

	Initial Eigenvalues			Rotation Sums of Squared Loadings		
Factor	Total	% of Variance	Cum. %	Total	% of Variance	Cum. %
1	5.55	25.23	25.23	4.49	20.39	20.39
2	2.06	9.37	34.60	3.13	14.21	34.60
3	1.54	6.98	41.57			
4	1.44	6.53	48.10			
5	1.35	6.14	54.24			
6	1.23	5.58	59.83			
7	1.09	4.95	64.77			
8	1.00	4.55	69.33			
9	.92	4.20	73.52			
10	.87	3.96	77.48			
11	.68	3.09	80.57			
12	.66	3.01	83.58			
13	.57	2.60	86.19			
14	.52	2.35	88.53			
15	.48	2.19	90.72			
16	.42	1.89	92.61			
17	.37	1.67	94.27			
18	.33	1.51	95.78			
19	.30	1.38	97.16			
20	.26	1.16	98.33			
21	.20	.88	99.21			
22	.17	.79	100.00			

his/her activity, wondering if the child was deaf, and staring or wandering with no purpose.

The second factor, which explained 14.21% of the total variance, included items relating to social interaction, joint attention, and receptive language. This factor was named Social Communication and included items 10, 14, 15, 17, 21, and 23 on the M-CHAT. These items addressed making eye contact, responding to his/her name, following another's point, looking at things that others are looking at, understanding what others are saying, and looking at a parent's face to check their reaction when faced with something unfamiliar.

Table 20

Rotated Component Matrix for Principal Components Analysis for Students with Autism

	Factors	
	1	2
Q1	.20	.25
Q2	**.56**	.29
Q3	.11	.20
Q4	**.41**	.38
Q5	**.75**	-.11
Q6	**.69**	.10
Q7	**.63**	.18
Q8	**.50**	.21
Q9	**.66**	.05
Q10	.22	**.55**
Q11	.27	.28
Q12	-.02	.37
Q13	**.61**	-.03
Q14	.06	**.77**
Q15	**.53**	**.54**
Q17	.17	**.60**
Q18	**.45**	.11
Q19	**.62**	.17
Q20	**.43**	.09
Q21	.07	**.61**
Q22	**.48**	.34
Q23	-.05	**.63**

Note. Factor loadings > .40 are in boldface.

While a two-factor solution was the best fit for the data, the two factors identified did not create discrete components for social communication skills and unusual behaviors as expected (see Table 20 for individual factor loadings). Social communication items contributed to both factors, while unusual behaviors contributed to only one factor.

Principal Components Analysis for Students without Autism

As discussed above, the Kaiser criterion, Cattell's scree test, and parallel analysis were again used to determine the number of factors to retain. A total of seven factors with eigenvalues greater than 1 were identified when the Kaiser criterion was applied, as shown in Table 21. As illustrated by Figure 7, Cattell's scree test indicated that two factors should be retained in the current analysis. Parallel analysis, also demonstrated by Table 21, suggested that two factors should be retained for the current analysis.

As there was agreement between parallel analysis and scree plot examination, principal components analysis limited to two factors using a varimax (orthogonal) rotation method was conducted for parent ratings on the M-CHAT for students without autism. As demonstrated in Table 22, results from this analysis indicated that a total of 38.11% of the variance was explained using a two-factor solution.

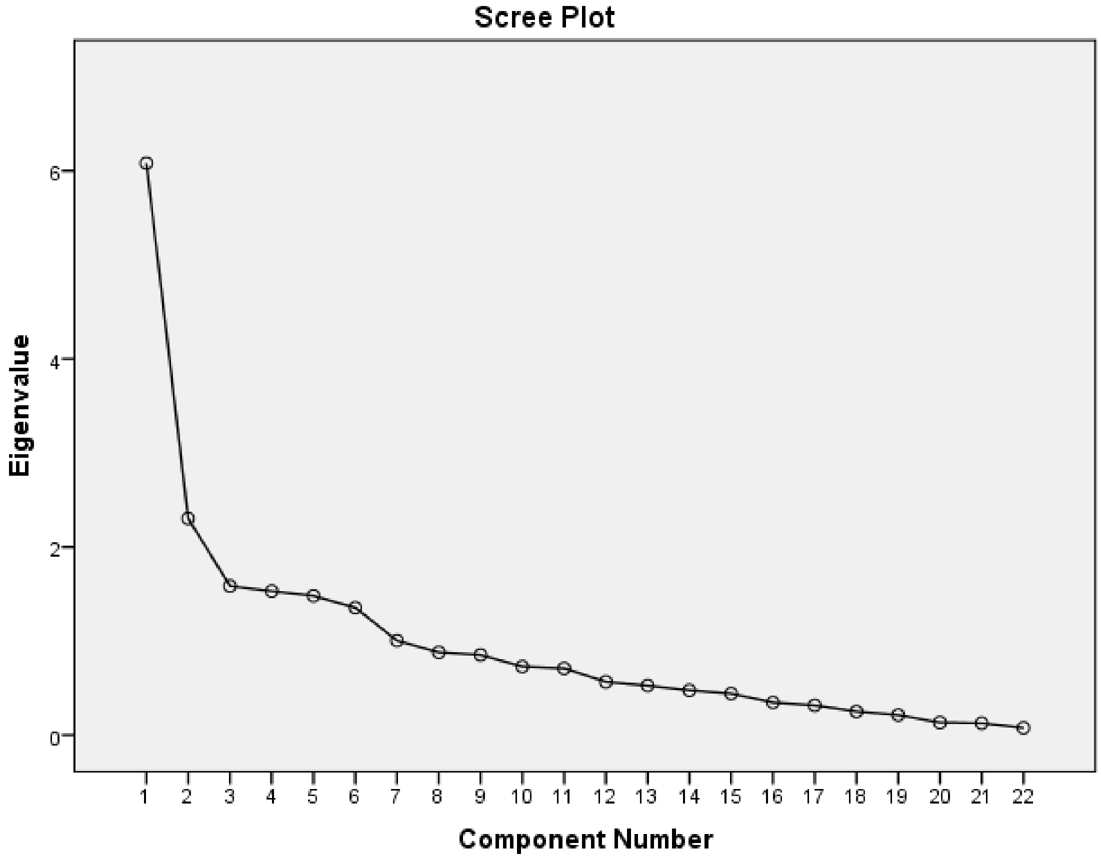

Figure 7. Scree plot of eigenvalues for students without autism indicating significant change of slope after two factors.

The first factor explained 27.64% of the total variance and included items related to play preferences, play skills, joint attention, and social interaction. This factor was named Play, Social Skills, and Joint Attention. Using the criterion of factor loadings greater than .40, this factor included items 3, 4, 5, 12, 13, 15, 17, 19, and 23 on the M-CHAT. These items addressed the following skills: Enjoying climbing on things such as stairs, enjoying playing games such as peek-a-boo or hide and seek, engaging in pretend play, smiling in response to others, engaging in imitation, following another's point, looking at things others are looking at, attracting a parent's attention to an activity, and checking a caregiver's face for their reaction when faced with something unfamiliar.

Table 21

Actual and Randomly Generated Eigenvalues Used for Parallel Analysis for Students without Autism

Factor	Actual	Parallel Analysis
1	6.08	1.90
2	2.31	1.73
3	*1.59*	*1.60*
4	1.53	1.50
5	1.48	1.41
6	1.36	1.32
7	1.01	1.25
8	.88	1.17
9	.85	1.10
10	.73	1.04
11	.71	.97
12	.57	.91
13	.53	.85
14	.48	.80
15	.44	.74
16	.35	.69
17	.32	.64
18	.25	.58
19	.21	.53
20	.13	.48
21	.13	.42
22	.08	.36

Note. Italics indicate the point at which actual eigenvalues were no longer greater than randomly generated eigenvalues. Random eigenvalues calculated using n = 113, number of variables = 22, replications = 1000.

Table 22

Total Variance Explained for Principal Components Analysis Using a Two-Factor Solution for Students without Autism

	Initial Eigenvalues			Rotation Sums of Squared Loadings		
Factor	Total	% of Variance	Cum. %	Total	% of Variance	Cum. %
1	6.08	27.64	27.64	6.08	27.64	27.64
2	2.31	10.48	38.11	2.31	10.48	38.11
3	1.59	7.21	45.32			
4	1.53	6.96	52.28			
5	1.48	6.74	59.02			
6	1.36	6.16	65.18			
7	1.01	4.57	69.75			
8	.88	4.00	73.75			
9	.85	3.88	77.63			
10	.73	3.31	80.94			
11	.71	3.23	84.17			
12	.57	2.58	86.74			
13	.53	2.40	89.14			
14	.48	2.17	91.31			
15	.44	2.01	93.32			
16	.35	1.58	94.90			
17	.32	1.44	96.34			
18	.25	1.14	97.49			
19	.21	.98	98.46			
20	.13	.61	99.07			
21	.13	.58	99.65			
22	.08	.36	100.00			

The second factor, which explained 10.48% of the total variance, included items relating to play preferences, social interaction, joint attention, and play skills. This factor was named Distal Pointing, Play Skills, and Social Interaction and included items 1, 2, 6, 7, 8, 15, 17, 19, 20, 21, and 22 on the M-CHAT. These items addressed enjoying being swung or bounced on the knee, taking an interest in other children, using a distal point to ask for things, using a distal point to indicate interest, engaging in functional play, following another's point, looking at things others are looking at, and attracting another's attention to an activity.

Table 23

Rotated Component Matrix for Principal Components Analysis for Students without Autism

	Factors	
	1	2
Q1	-.10	**.67**
Q2	.01	**.59**
Q3	**.47**	.02
Q4	**.43**	.16
Q5	**.73**	.01
Q6	.17	**.64**
Q7	.10	**.62**
Q8	.38	**.41**
Q9	**.76**	.20
Q10	.22	.33
Q11	-.21	.19
Q12	**.76**	.09
Q13	**.82**	-.10
Q14	.33	.27
Q15	**.52**	**.64**
Q17	**.54**	**.55**
Q18	.22	.31
Q19	**.54**	**.46**
Q20	.02	**.46**
Q21	.32	**.68**
Q22	.01	**.49**
Q23	**.42**	.24

Note. Factor loadings > .40 are in boldface.

Similar to above, while a two-factor solution was the best fit for the data, overall the two factors identified did not create discrete components for social communication skills and unusual behaviors as expected (see Table 23 for individual factor loadings). The factors identified were very similar in nature, each with varying aspects of social interaction, play, distal pointing, and joint attention. Unusual behaviors and communication did not significantly contribute to either factor.

Conclusions of Principal Components Analyses

As discussed above, dissimilar items on the M-CHAT loaded onto the factors identified for students with and without autism. Two factors were identified for each group but were different in nature. While factors for both groups contained varying amounts of social interaction, play skills and joint attention, no unusual behaviors loaded onto either component for students without autism, while unusual behaviors loaded strongly onto the first factor for students with autism. Therefore, the groups were not combined for a single principal components analysis.

It was hypothesized that the M-CHAT would not measure separate factors for social and communication deficits, based on factor analysis, cluster analysis, and principal components analysis conducted on other autism evaluation instruments (Constantino et al., 2004; Goldstein &

Naglieri, 2010; Gotham, Risi, Pickles, & Lord, 2007; Robertson, Tanguay, L'Ecuyer, Sims, & Waltrip, 1999). Overall, the hypothesis was accepted. There was not sufficient evidence to suggest separate factors for social vs. communication skills, neither for students with autism, nor students without autism. However, while two-factor solutions were the best fit for both groups, neither groups' factors clearly discriminated between social communication and unusual behaviors. Overall, the M-CHAT is not measuring autism as per the newly recommended model. For students with autism the M-CHAT is measuring two components, one containing most core features of autism, and one containing only social communication deficits. For students without autism the M-CHAT appears to be measuring a variety of social and plays skills along with joint attention, with very little contribution from communication or unusual behavior items.

Summary

This chapter addressed the findings of the current research study examining the use of the M-CHAT as an autism screening tool for preschool-age special education students. Overall, findings indicated that the M-CHAT correctly classified 62% of students in the current sample, while incorrectly classifying 38%. Sensitivity was .64 for the current sample while the specificity was .60. The PPV for the M-CHAT was .61 and the NPV was .64. This held true regardless of the scoring method. No differences were found with the use of traditional scoring vs. Best7 scoring methods. It was found that the failure of Question 13 ("Does your child imitate you? [e.g., you make a face - will your child imitate it?]"; Robins, Fein, & Barton, 1999), and Question 2 ("Does your child take an interest in other children?") were the best predictors of students receiving an educational classification of autism, though no new scoring methods could be introduced for the M-CHAT.

Results of principal components analysis indicated that the M-CHAT is composed of two components, though the emphasis of these components was different for students with and without autism. While the factors were sufficiently different to prevent combining the groups for a single factor analysis, neither group demonstrated a clear delineation between social communication and unusual behaviors.

CHAPTER V

DISCUSSION

The purpose of this research study was to examine the utility of the Modified Checklist for Autism in Toddlers (M-CHAT; Robins, Fein, & Barton, 1999) in a preschool-age special education sample, using both traditional scoring methods and the newly introduced Best7 scoring procedure. This study also examined which items on the M-CHAT were most associated with students' receiving an educational classification of autism, as discriminated from students with other developmental delays. In addition, this study attempted to determine whether a two-factor model of autism identification (i.e., social communication deficits and unusual/repetitive behaviors) or a different model best described student performance on the M-CHAT.

To complete this research, archival data were gathered for 266 participants in the Berks County Intermediate Unit (BCIU) Early Intervention (EI) program (n = 133 students with autism and n = 133 students with other developmental delays), including M-CHAT results and demographic information (age, sex, race, and educational classification). Complete data sets were available for 222 of the original 266 students. Descriptive statistics; calculations of sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV); ROC analysis; stepwise discriminant function analysis; and principal components analysis were used to answer the following research questions.

Discussion of Findings

Research Question 1

The first research question was: How does the M-CHAT perform as an autism screening measure in the current sample of preschool-age students (30-72 months) based on the traditional scoring procedures established for 16-30 month olds? It was hypothesized that there would be a high rate of false positives (or low PPV; number of students who failed the M-CHAT but do not have autism), as indicated by previous research (Robins, 2008; Robins, Fein, Barton, & Green, 2001). It was also hypothesized that the sensitivity of the M-CHAT would be lower than reported in the initial validation study, as research has suggested that the M-CHAT is less sensitive in older children (Eaves, Wingert, & Ho, 2006; Snow & Lecavalier, 2008).

Results of the current study indicate that this hypothesis was confirmed. A large number of false positives (students who failed the M-CHAT but do not have autism) were identified. Overall, 45 of 113 students who did not have autism failed the M-CHAT, or 39.8% of all students who did not have autism. Additionally, the sensitivity of the M-CHAT in the current study (.64) was much lower than the range of .87 - .97 that was reported in the initial validation study (Robins et al., 2001).

These results are consistent with the findings of Eaves et al. (2006), where it was noted that a substantial number of false positives were found when the M-CHAT was used as a screening tool for 2 and 3 year olds. Other follow-up studies examining the M-CHAT with older populations also reported lower sensitivity and specificity than what has been found for 16-30 month olds (Eaves et al., 2006; Snow & Lecavalier, 2008). These studies also examined the M-CHAT's use with Level II, or high-risk populations, which is likely an additional cause for the poor predictive validity found in the current sample. Glascoe (2005) suggested that acceptable levels of sensitivity and specificity for screening measures should be 70%, when screening measures are used as in primary care settings, or as Level I screeners. When screening measures are being used in tertiary settings, or as a Level II screening instrument as in the current study, the argument has been made that these levels should be even higher "in order to produce maximum correct classification" (Snow & Lecavalier, 2008, p. 641). None of the calculated levels of sensitivity (.64), specificity (.60), PPV (.61), or NPV (.64) in the current sample approached the minimum accuracy threshold of 70% for a Level I screener, let alone demonstrated even higher levels of accuracy expected for Level II screeners.

While ROC analysis indicated significant findings (in that there was a relationship between failure of M-CHAT questions and educational classification that was not due to chance), the M-CHAT overall correctly classified only 62% of students. As the expected outcome for random chance (e.g., flipping a coin) would be accurate approximately 50% of the time, the M-CHAT was only slightly better than random chance. Significance aside, this is a poor outcome for an autism screening measure.

Research Question 2

The second research question was: How does the M-CHAT perform as an autism screening measure in the current sample of preschool-age students (30-72 months) based on the Best7 scoring procedures suggested for 16-30 month olds? It was hypothesized that there would be a more moderate rate of false positives (number of students who failed the M-CHAT but do not have autism), as the Best7 criteria were suggested in order to decrease the high rate of false positives previously reported with traditional scoring. It was also hypothesized that sensitivity of the M-CHAT would again be lower than reported in the initial validation of the study, due to the age range and nature of the current sample.

Results indicated that the hypothesis of generating fewer false positives with the Best7 scoring method was rejected. A similarly large number of false positives were identified with the Best7 scoring method as were identified with the traditional scoring method. Overall, 45 of 113, or 39.8% of students who did not have autism failed the M-CHAT using the Best7 scoring method. The hypothesis that the sensitivity of the M-CHAT would again be lower than in the initial validation study, however, was confirmed. The sensitivity of the M-CHAT in the current study (.64) was much lower than the range of .87 - .97 reported in the initial validation study (Robins et al., 2001).

As discussed above, research examining the use of the M-CHAT for older students has consistently found lower sensitivity, specificity, PPV, and NPV than when it is used for 16-30 month olds (Eaves et al., 2006; Snow & Lecavalier, 2008). The current study supports these findings. Beyond the initial report presented at the International Meeting for Autism Research (Robins et al., 2010), there has been no published research regarding the

Best7 scoring method; however, the findings of this study do not support the claim that this method results in lowering the number of false positives, at least not for preschool-age students.

Research Question 3

The third research question was: What are the differences in the M-CHAT's performance as an autism screening measure in the current sample, based on the use of the traditional scoring procedures vs. the Best7 scoring procedures? It was hypothesized that fewer false positives would be identified with the use of the Best7 scoring procedures, while the sensitivity would remain relatively constant; as the intention of the Best7 scoring algorithm was to increase PPV without significantly affecting sensitivity and specificity (Robins et al., 2010). This hypothesis was rejected, as no differences were found in the rate of false positives or in the sensitivity of the M-CHAT when the two scoring methods were compared.

As discussed under Research Question 2, there has been no published research regarding the use of the Best7 scoring procedures; however the current study does not show any differences between traditional methods and Best7 methods when used with preschool-age students. Further, when the M-CHAT forms were scored for current study it was noted that none of the children who failed or passed the M-CHAT did so on the basis of critical items for either scoring method. All children passed or failed based on the criteria of failing three or more total items on the M-CHAT. Therefore, the use of critical items (either the traditional six or the Best7) does not seem to improve scoring outcomes for preschool-age children. Snow and Lecavalier (2008) found similar results, where the cutoff of failing three items was more accurate than using the cutoff criterion of failing two critical items.

Research Question 4

The fourth research question was: Which items on the M-CHAT are most predictive of preschool-age students (30-72 months) in the current sample receiving an educational classification of autism? Robins et al. (2001) originally identified critical items as children's ability to take an interest in other children, use their index finger to point to objects of interest, bring objects to their parents to show them something, engage in imitation, respond to their name, and follow a point in order to look at something across the room). When the Best7 scoring algorithm was proposed, the critical item of imitation was eliminated and two critical items were added, including engaging in pretend play and wondering if the child was deaf (Robins et al., 2010). It was hypothesized that similar items would be identified as strong predictors of autism; however, given the age of the sample it was also hypothesized that more items than these six to seven would be identified as strongly predictive of autism.

The hypothesis was partially accepted, in that similar items were found to be highly predictive of autism; however, fewer items, rather than more, were found to be predictive of autism in the current model. Both items identified in the current analysis (ability to imitate and taking an interest in other children) were originally identified as being highly predictive of autism in the traditional scoring approach. On the other hand, only the item related to taking an interest in other children was identified as being predictive of autism using the Best7 scoring method. In addition, only two items were found to be significantly predictive of autism using the current sample, as opposed to six and seven items previously identified using the traditional and Best7 scoring approaches, respectively.

The purpose of this research question was to identify whether a better scoring method could be introduced for older students, based on newly identified critical items. However, all students who failed via the traditional scoring method would have failed using the two items identified in this analysis, as both items had already been identified as critical. This would not have improved the predictive validity of the M-CHAT in any meaningful way. These findings support the use of traditional scoring as opposed to the Best7 scoring in regards to critical items for preschool-age students. However, it was previously noted that none of the children who passed or failed the M-CHAT in the current study did so on the basis of critical items, rendering this comparison meaningless.

It is presumed that fewer items were identified as being strongly predictive of autism, rather than more as was hypothesized, due to the age of the students in the current sample. As children age, they are likely to show a wider range of characteristics relating to autism, hence the basis for the original hypothesis (Coonrod & Stone, 2005; Robertson, Tanguay, L'Ecuyer, Sims, & Waltrip, 1999; Snow & Lecavalier, 2008). However, it is likely that as characteristics of autism became more heterogeneous with age, each individual item on the M-CHAT decreased in predictive value as the number of identifiable characteristics increased.

Attempts to improve the predictive validity of the M-CHAT by changing the cutoff criteria in the current study were also not successful. If the cutoff criterion was changed to failing any two items, rather than three, an additional six students with autism could be correctly classified, raising the sensitivity to .70. However, this meant that an additional 13 students who did not have autism would also have failed the screening, which reduced the PPV to .57. This created an unsatisfactory amount of false positives, which was already at an unacceptably high level with the current scoring method. Changing the cutoff criteria to failing one item was also not possible. While an additional 26 students with autism would be correctly classified, all but 30 of the 113 students without autism would have failed the screener, rendering it essentially useless. Overall, the M-CHAT does not appear to be an exemplary autism screening tool for preschool-age students.

Research Question 5

The fifth research question was: Does a principal components analysis of parent ratings indicate that the M-CHAT is measuring two components of autism (i.e., social communication deficits and unusual/repetitive behaviors), as per the newly recommended autism identification model, or will a different solution best represent findings from the current sample? It was hypothesized that the M-CHAT would not measure separate factors for social and communication deficits, based on factor analysis, cluster analysis, and principal components analysis conducted on other autism evaluation instruments (Constantino et al., 2004; Goldstein & Naglieri, 2010; Gotham, Risi, Pickles, & Lord, 2007; Robertson, Tanguay, L'Ecuyer, Sims, & Waltrip, 1999).

Overall, the hypothesis was accepted. There was not sufficient evidence to suggest separate factors for social vs. communication skills, neither for students with autism, nor students without autism. Yet, while two-factor solutions were the best fit for both groups, neither groups' factors clearly discriminated between social communication and unusual behaviors. Overall, the M-CHAT is not measuring autism as per the newly recommended model. For students with autism the M-CHAT is measuring two components, one containing most core features of autism, and one containing only social communication deficits. For students without autism the M-CHAT appears to be measuring a variety of social and plays skills along with joint attention, with very little contribution from communication or unusual behavior items.

The major difference was that for students with disabilities other than autism, the factors did not include communication or unusual behaviors. Overall, the factors for these students included varying aspects of social interaction, distal pointing, play skills, and joint attention. For students with autism, communication and unusual behaviors were a critical part of the components identified. This may explain in part why there were so many false positives in the current sample. Due to the small number of items addressing unusual behaviors (three) and the fact that the majority of items related to social interaction and communication do not require the deficits to be stereotypic in any way, students demonstrating developmental delays could easily fail three or more items without any of those items clearly discriminating between autism and other delays.

For students with autism, the findings in the current study were very similar to the cluster analysis conducted on the Autism Diagnostic Interview, Revised (ADI-R; Constantino et al., 2004). In that study the authors found two clusters, one encompassing all areas traditionally associated with autism, and one cluster including social deficits and communication. A follow-up principal components analysis conducted on the ADI-R, however, did not indicate the same findings, as one main factor encompassing all deficits associated with autism and 13 smaller factors were identified. Factor analysis of the

Social Responsiveness Scale (SRS; Constantino & Gruber, 2005), yielded similar findings to the principal components analysis of the ADI-R, with one main factor encompassing all domains typically associated with autism and multiple smaller, overlapping factors. These findings were not consistent with the principal components analysis of the M-CHAT.

The findings of the current study were also dissimilar to the factor analysis studies examining diagnostic instruments for autism, such as the Autism Diagnostic Observation Schedule (Gotham, Risi, Pickles, & Lord, 2007) and the Autism Spectrum Rating Scales (Goldstein & Naglieri, 2010). In these studies, designed to examine the underlying factor structure of autism, a two-factor model of autism (social communication and unusual behaviors) was found to explain the most variance and led to improved predictive validity. This two-factor model is now considered best practice for identifying autism in the mental health field (APA, 2012, 2013). Overall, the M-CHAT is not measuring autism according to this model, as its components did not break down into a simple two-factor model of social communication and unusual behaviors.

All of these studies, despite the type of analysis conducted, did have one conclusion in common: Separate factors for social and communication domains were not supported. The current research study supports these findings, as there was no evidence to support separate social and communication factors, although the factors did not break down into a simple two-factor solution as initially expected.

Limitations

As discussed in Chapter 1, a major assumption of the current study was that participants had been identified under the correct educational classification. It is possible, however, that some students with autism may not have been correctly identified, especially given that the current study questioned whether the M-CHAT is an appropriate autism screening measure for preschool-age students. If the M-CHAT was not accurately recognizing students in need of further testing for autism there may have been a small number of students who were classified incorrectly as having a developmental delay only. There is no way to know with certainty whether any students were missed, although it is clear that the M-CHAT only correctly identified 64% of children who had been previously diagnosed with autism. As the M-CHAT does not appear to be an accurate autism screening tool for the current sample, any missed cases would only worsen its predictive validity, not strengthen it. This limitation does not significantly affect the findings of the current study.

Additionally, there is a certain amount of limitation inherent when examining autism screeners in special education populations, due to significant variability between states. The disability category of developmental delay, for example, is left for states to determine (IDEIA, 2004, §300.111). While Pennsylvania defines developmental delay as either a 25% delay based on chronological age or a score of 1.5 standard deviations below the mean (Pennsylvania Chapter 14 Special Education Regulations, 2008, §14.153), other states have varying definitions, indicating that children who qualify for early intervention services in one state may not qualify in another. Even the educational definition of autism, while clearly defined in the IDEIA federal regulations (§300.8), is subject to a certain amount of variability between states. Colorado, for example, has chosen to include autism under the category of other health impairment, which was changed to physical disability in their state regulations, rather than using autism as a stand-alone category. Colorado also requires a medical diagnosis for students to be identified as having a physical disability (Noland & Gabriels, 2004). Four other states have developed different definitions of autism that do not precisely follow federal guidelines, and one state stipulates that pervasive developmental disorder, not otherwise specified (PDD-NOS) should not be included under the state's educational definition of autism (Muller & Markowitz, 2004), contrary to best practice recommendations (Ikeda, 2002; Shriver, Allen, & Matthews, 1999). In addition, some states require that the diagnosis of autism must be made by a specific professional, including a physician (11 states), a licensed psychologist (nine states), a psychiatrist (eight states), a school psychologist (seven states, including Pennsylvania), and a developmental pediatrician (one state; Muller & Markowitz, 2004). This variability results in somewhat vague models of autism identification at the national level (Ikeda, 2002) along with differing approaches to evaluating students suspected of having an autism spectrum disorder, and limits educational research to the state level.

In consequence, this research study took place at a single intermediate unit in Pennsylvania, one of the states that requires a school psychologist when evaluating for autism (Pennsylvania Chapter 14 Special Education Regulations, 2008, §14.123). While the sample represented a diverse population, the findings should be interpreted with caution. There are currently no other published studies examining the use of the M-CHAT with special education samples. As this is potentially the first study of its kind, more information is needed to ensure that these findings hold true across various states, samples, and demographics.

Implications for Future Research

Based on the limitations described above, an apparent recommendation is that research on the M-CHAT's predictive validity be conducted in a variety of educational settings and across different states to determine whether outcomes continue to be poor and to further validate the findings of the current study. However, it is expected that additional research will continue to demonstrate low sensitivity and PPV for the M-CHAT, given that other studies examining the M-CHAT's use in populations older than 16-30 months and with Level II populations also question its predictive validity (Eaves et al., 2006; Snow & Lecavalier, 2008), regardless of the location or setting where it was used. Research may be better served concentrating on improving the M-CHAT with a focus on preschool-age students and high-risk populations. Possible improvements to the M-CHAT may include adding questions relating to expressive language and/or atypical patterns of language. The M-CHAT currently addresses only receptive and nonverbal language. This was likely done due to the young age range it was intended to screen, but adding these types of questions may help improve predictive validity for older students. It was also noted that there are only three items related to unusual behaviors on the M-CHAT. Increasing the number of these items, along with tying some of the communication and social interaction items to atypical patterns of development rather than just delays, may help improve discrimination between students with autism and other disabilities. In addition, the use of critical items as a scoring approach was not supported by the results of this study. As researchers attempt to improve current M-CHAT questions, focus should be given to identifying new items that accurately differentiate between students with autism and those without, or perhaps the idea of critical items should be abandoned altogether in favor of a new cutoff score that can demonstrate adequate sensitivity, specificity, PPV, and NPV in future studies. Further, the M-CHAT was not found to include clearly defined factors addressing social communication and unusual behaviors for preschool-aged students, based on principal components analyses completed as part of this study. Results suggested that the M-CHAT does encompass two components, but these do not clearly demarcate social communication deficits and unusual behaviors, possibly indicating that the current questions are not appropriately assessing autism characteristics for preschool-age children. As studies suggest that autism evaluation instruments have better predictive validity when based on these simplified factors (Goldstein & Naglieri, 2010; Gotham et al., 2008; Gotham, Risi, Pickles, & Lord, 2007; Oosterling et al., 2010), attempts to improve the M-CHAT and/or to create new autism screening instruments should focus on the inclusion of items that clearly address these two factors in a well-defined manner.

Given the questionable predictive validity of the M-CHAT for preschool-age students, another direction for research would be to focus on creating new accurate, reliable autism screening measures for preschool-aged students. This could occur through a combination of creating new screening instruments, as well as improving the current questions on the M-CHAT to better assess preschool-age students, as described above. As discussed in Chapter 2, there is a gap in the available screeners for the preschool-age population. Available Level II screeners tend to focus on children ages 4 and older, and additionally could be cost-prohibitive if they were to be administered to all children referred for early intervention (EI) services. Similar problems are noted with Level I screeners. Of the Level I screeners that had the advantage of being free, they tended to focus on very young children, often as young as 12-24 months. There is a lack of widely available screening instruments that can be universally administered to preschool-age students who are referred to EI services in order to make accurate referrals for autism evaluations.

Implications for School Psychology

Autism is a growing disability, with numbers now approaching 1 in 50 school-age children (Blumberg et al., 2013), and school psychologists should be prepared to accurately assess and identify these children (Wilkinson, 2011). Accurate means of screening for and identifying autism are necessary to ensure that these children receive appropriate services. Early identification is vital, as research consistently indicates that when interventions are provided at a young age, outcomes are improved (Crane & Winsler, 2008; Levy et al., 2007; National Research Council, 2001; Robins & Dumont-Mathieu, 2006).

As hitherto discussed, there are a number of autism screening tools for school-age students and for very young children, but there are few measures, if any, that focus on preschool-age children that can be administered to large at-risk populations. This is surprising, given that most children with autism are identified at this age (Shattuck et al., 2009). Until more reliable and valid screening tools have been developed and are widely available, school psychologists must remain aware of the limitations of the autism screening tools they currently employ for young children. Results of these screening measures, especially the M-CHAT, should be interpreted with caution given the findings of this study. Passage or failure of the M-CHAT should never be the only, or even the primary, tool for making decisions as to whether a child is in need of evaluation for autism. Multiple sources of data should be used when making decisions regarding diagnostic testing.

A national model for autism identification in educational settings would be a step in the right direction, though this is complicated by the variability between state requirements. At this time, school psychologists and other school professionals involved in the diagnostic process must continue to follow state-mandated procedures and recommended models of autism identification. In Pennsylvania this encompasses a stepwise model using multiple sources of data including observations, an evaluation of current educational performance and levels of functioning, parent interview and developmental history, autism-related checklists and/or Level II screening measures, and the use of "gold-standard" evaluation instruments such as the Autism Diagnostic Observation Schedule, Second Edition (Lord, et al., 2012; Levy et al., 2007). In addition, evaluators must possess an in-depth knowledge of any cultural considerations for the population with whom they practice (Overton, Fielding, & Garcia de Alba, 2007) and how these may affect diagnostic decision making. Understanding the recommended models for autism identification and having a basic familiarity with evaluation instruments, however, is not enough for a school psychologist, or any practitioner, to identify a child with an autism spectrum disorder. Experts agree that the diagnosis of autism should be made only by those who have specialized training, significant clinical experience, and extensive expertise with autism (Levy et al., 2007; Lord & Corsello, 2005; Schwartz & Davis, 2008).

Conclusions

In conclusion, the M-CHAT was not found to be an effective autism screening tool for preschool-age special education students. The current study indicates that the M-CHAT correctly classified only 62% of students, which is little better than what could be expected from random chance. Attempts to improve scoring methods or identify particular questions which may improve the M-CHAT's discriminative ability were unsuccessful. While further studies are needed to validate these findings with additional special education samples, overall, it cannot be recommended that the M-CHAT be used as the primary autism screening method for preschool-age students.

It should be noted that the M-CHAT is an appropriate screening instrument *when used for the age range and purpose for which it was created*. The difficulty is inherent when it is not used for its intended purpose. When the age range is increased to include preschool-age students and the instrument is used with Level II populations, the M-CHAT becomes far less accurate for diagnostic decision making. There is a clear need for improved screening measures for preschool-age students.

Summary

The number of young children with autism is ever-increasing, and school psychologists are more frequently required to identify these students (Blumberg et al., 2013; Wilkinson, 2011). Valid screening tools are needed in order to focus school psychologists' time on those students in need of intensive evaluations in EI programs.

The purpose of this study was to examine the utility of the M-CHAT in a preschool-age special education sample, using both traditional scoring methods and the newly introduced Best7 scoring procedure. This study also examined which items on the M-CHAT were most associated with students' receiving an educational classification of autism, as discriminated from students with other developmental delays. In addition, this study examined whether a two-factor model of autism identification (i.e., social communication deficits and unusual behaviors) or a different model best described student performance on the M-CHAT.

Overall, findings indicated that the M-CHAT correctly classified 62% of students in the current sample, while incorrectly classifying 38%. Sensitivity was .64 for the current sample while the specificity was .60. The PPV for the M-CHAT was .61 and the NPV was .64. No differences were found with the use of traditional scoring vs. Best7 scoring methods. Results of stepwise discriminant function analysis indicated that failure of Question 13 (Does your child imitate you?), and Question 2 (Does your child take an interest in other children?) were the best predictors of students receiving an educational classification of autism; however this did not lead to an improved scoring method for the current sample. Attempts to improve the M-CHAT's performance by changing the cutoff criteria were unsuccessful.

Results of principal components analyses indicated that the M-CHAT is composed of two components, though the emphasis of these components was different for students with and without autism. While the factors were sufficiently different to prevent combining the groups for a single factor analysis, neither group demonstrated a clear delineation between social communication and unusual behaviors. Thus, the M-CHAT does not appear to be measuring autism as per the newly recommended two-factor model of social communication and unusual behaviors.

Overall, the M-CHAT was not found to be an accurate autism screening tool in the current sample, consistent with previous research examining its use in older populations. Future research should focus on improving current M-CHAT questions to better assess preschool-age students and/or developing new autism screening measures designed specifically for preschool-age students. Improvements to M-CHAT for older students may include the addition of questions specifically addressing verbal communication as opposed to only nonverbal and receptive language, as well as ascertaining that the M-CHAT's factor structure supports current autism identification recommendations of social communication and unusual behaviors factors. Until such screening tools are available, school psychologists must remain aware of the limitations of the autism screening tools and evaluation instruments they currently employ and should interpret results with caution. Efforts must be made to ensure that practitioners are experienced with autism identification and have a thorough understanding of the variety of ways in which this disability may manifest itself in preschool-age children. Accurate identification is crucial to ensuring that children gain access to much needed services and interventions at an early age.

References

Akshoomoff, N., Corsello, C., & Schmidt, H. (2007). The role of the Autism Diagnostic Observation Schedule in the assessment of autism spectrum disorders in school and community settings. *California School Psychologist, 11*, 7-19.

Allison, C., Baron-Cohen, S., Wheelwright, S., Charman, T., Richler, J., Pasco, G., & Brayne, C. (2008). The Q-CHAT (Quantitative CHecklist for Autism in Toddlers): A normally distributed quantitative measure of autistic traits at 18–24 months of age: Preliminary report. *Journal of Autism and Developmental Disorders, 38*, 1414–1425. doi:10.1007/s10803-007-0509-7

American Psychiatric Association. (1952). *Diagnostic and statistical manual: Mental disorders*. Washington, DC: American Psychiatric Association: Mental Hospital Service.

American Psychiatric Association. (1968). *Diagnostic and statistical manual of mental disorders* (2nd ed.). Washington, DC: Author.

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, DC: Author.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

American Psychiatric Association. (2012). *DSM-5 development, 299.00 autism spectrum disorder*. Retrieved from http://www.dsm5.org

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Baio, J. (2012). *Prevalence of autism spectrum Disorders-Autism and Developmental Disabilities Monitoring Network, 14 sites, United States, 2008: Surveillance summaries.* Retrieved from Centers for Disease Control and Prevention website: http://www.cdc.gov/mmwr/preview/mmwrhtml/ss6103a1.htm?s_cid=ss6103a1_w

Baird, G., Charman, T., Baron-Cohen, S., Cox, A., Swettenham, J., Wheelwright, S., & Drew, A. (2000). A screening instrument for autism at 18 months of age: A 6 year follow-up study. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*, 694-702.

Baron-Cohen, S., Allen, J., & Gillberg, C. (1992). Can autism be detected at 18 months? The needle, the haystack, and the CHAT. *British Journal of Psychiatry, 161,* 839-843.

Baron-Cohen, S., Cox, A., Baird, G., Swettenham, J., Nightengale, N., Morgan, K., ... Charman, T. (1996). Psychological markers in the detection of autism in infancy in a large population. *British Journal of Psychiatry, 168,* 158-163.

Baron-Cohen, S., Wheelwright, S., Cox, A., Baird, G., Charman, T., Swettenham, J., ... Doehring, P. (2000). The early identification of autism: The CHecklist for Autism in Toddlers (CHAT). *Journal of the Royal Society of Medicine, 93*, 521-525.

Blumberg, S. J., Bramlett, M. D., Kogan, M. D., Schieve, L. A., Jones, J. R., & Lu, M. C. (2013). *Changes in prevalence of parent-reported autism spectrum disorder in school-aged U.S. children: 2007 to 2011–2012.* Retrieved from Centers for Disease Control and

Prevention website: http://www.cdc.gov/nchs/data/nhsr/nhsr065.pdf

Brudereck, J. (2012, September 20). Reading sheds 'poorest' rank, but poverty still a challenge. *Reading Eagle.* Retrieved from http://readingeagle.com

Carter, A. S., Davis, N. O., Klin, A., & Volkmar, F. R. (2005). Social development in autism. In F. R. Volkmar, R. Paul, A. Klin, & D. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders* (3rd ed., pp. 312-334). Hoboken, NJ: Wiley.

Cattell, R. B. (1966). The Scree test for the number of factors. *Multivariate Behavioral Research, 1*, 245-276.

Chapter 14 Special Education Services and Programs State Regulations. (2008). Commonwealth of Pennsylvania. Retrieved from http://pattan.net-website.s3.amazonaws.com/images/file/2011/08/15/sidebyside021209.pdf

Constantino, J. N. (2012). *Social Responsiveness Scale, second edition*. Los Angeles, CA: Western Psychological Services.

Constantino, J. N., & Gruber, C. P. (2005). *Social Responsiveness Scale*. Los Angeles, CA: Western Psychological Services.

Constantino, J. N., Gruber, C. P., Davis, S., Hayes, S., Passanante, N., & Przybeck, T. (2004). The factor structure of autistic traits. *Journal of Child Psychology and Psychiatry, 45*, 719–726.

Coonrod, E. E., & Stone, W. L. (2005). Screening for autism in young children. In F. R. Volkmar, R. Paul, A. Klin, & D. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders* (3rd ed., pp. 707–729). Hoboken, NJ: Wiley.

Corsello, C., Hus, V., Pickles, A., Risi, S., Cook, E. H., Leventhal, B. L., & Lord, C. (2007) Between a ROC and a hard place: Decision making and making decisions about using the SCQ. *Journal of Child Psychology and Psychiatry, 48*, 932–940.

Crane, J. L., & Winsler, A. (2008). Early autism detection: implications for pediatric practice and public policy. *Journal of Disability Policy Studies, 18*, 245-253. doi:10.1177/1044207307311527

Dumont-Mathieu, T. & Fein, D. (2005). Screening for autism in young children: The Modified Checklist for Autism in Toddlers (M-CHAT) and other measures. *Mental Retardation and Developmental Disabilities Research Reviews, 11*, 253-262. doi:10.1002/mrdd.20072

Eaves, L. C., Wingert, H., & Ho, H. (2006). Screening for autism: Agreement with diagnosis. *Autism, 10*, 229-242. doi:10.1177/1362361306063288

Factor, D. C., Freeman, N. L., & Kardash, A. (1989). Brief report: A comparison of DSM-III and DSM-III-R criteria for autism. *Journal of Autism and Developmental Disorders, 19*, 637-640.

Filipek, P. A., Accardo, P. J., Ashwal, S., Baranek, G. T., Cook, E. H., Dawson, G., ... Volkmar, F. R. (2000). Practice parameter: Screening and diagnosis of autism: Report of the quality standards subcommittee of the American Academy of Neurology and the Child Neurology Society. *Neurology, 55*, 468-479. Retrieved from http://www.childneurologysociety.org/

Filipek, P. A., Accardo, P. J., Baranek, G. T., Cook, E. H., Dawson, G., Gordon, B., ... Volkmar, F. R. (1999). The screening and diagnosis of autistic spectrum disorders. *Journal of Autism and Developmental Disorders, 29*, 439-484.

Frith, U. (2003). *Autism: Explaining the enigma* (2nd ed.). Oxford, UK: Blackwell.

Frombonne, E. (2005). Epidemiological studies of pervasive developmental disorders. In F. R. Volkmar, R. Paul, A. Klin, & D. Cohen (Eds.), *Handbook of autism and* *pervasive developmental disorders* (3rd ed., pp. 42-69). Hoboken, NJ: Wiley.

Gillberg, C. (1998). Asperger syndrome and high-functioning autism. *British Journal of Psychiatry, 172*, 200–209.

Gilliam, J. E. (2004). *Gilliam Autism Rating Scale, second edition.* Austin, TX: Pro-Ed.

Glascoe, F. P. (2000). Early detection of developmental and behavioral problems. *Pediatrics in Review, 21*, 272-280.

Glascoe, F. P. (2005). Screening for developmental and behavioral problems. *Mental Retardation and Developmental Disabilities Research Reviews, 11*, 173–179.

Goldstein, S., & Naglieri, J. A. (2010). *Autism Spectrum Rating Scales technical manual.* North Tonawanda, NY: Multi-Health Systems.

Goldstein, S., & Naglieri, J. A. (2011). *Autism: New data, new ideas, and the Autism Spectrum Rating Scale*. Workshop presented at the National Association of School Psychologists Annual Convention, San Francisco, CA.

Gotham, K., Risi, S., Dawson, G., Tiger-Flusberg, H., Joseph, R., Carter, A., ... Lord, C. (2008). A replication of the Autism Diagnostic Observation

Schedule (ADOS) revised algorithms. *Journal of the American Academy of Child and Adolescent Psychiatry, 47*, 642-651. Retrieved from http://www.jaacap.com/

Gotham, K., Risi, S., Pickles, A., & Lord, C. (2007). The Autism Diagnostic Observation Schedule: Revised algorithms for improved diagnostic validity. *Journal of Autism and Developmental Disorders, 37*, 613-627. doi:10.1007/s10803-006-0280-1

Hanley, J. A., & McNeil, B. J. (1983). A method of comparing the areas under receiver operating characteristics curves derived from the same cases. *Radiology, 148*, 839-843. Retrieved from http://www.medicine.mcgill.ca/epidemiology/hanley/software/Hanley_McNeil_Radiology_83.pdf

Hippler, K., & Klicpera, C. (2003). A retrospective analysis of the clinical case records of 'autistic psychopaths' diagnosed by Hans Asperger and his team at the University Children's Hospital, Vienna. *Philosophical Transactions of the Royal Society: Biological Sciences, 358*, 291–301. doi:10.1098/rstb.2002.1197

Ikeda, M. (2002). Best practices for supporting students with autism. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (pp. 1501-1512).

Bethesda, MD: National Association of School Psychologists.

Individuals with Disabilities Education Act. (1990). Federal Regulations. Retrieved from http://www.ada.gov/archive/adastat91.htm

Individuals with Disabilities Education Improvement Act. (2004). Federal Regulations Part 300. Retrieved from http://pattan.net-website.s3.amazonaws.com/images/file/2011/08/15/sidebyside021209.pdf

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child, 2,* 217-250. Retrieved from http://www.aspires-relationships.com

Kleinman, J. M., Robins, D. L., Ventola, P. E., Pandey, J., Boorstein, H. C., Esser, E. L., ... Fein, D. (2008). The Modified Checklist for Autism in Toddlers: A follow-up study investigating the early detection of autism spectrum disorders. *Journal of Autism and Developmental Disorders, 38*, 827-839. doi:10.1007/s10803-007-0450-9

Kolvin, I. (1971). Studies in the childhood psychoses: I. Diagnostic criteria and classification. *British Journal of Psychiatry, 118*, 381-384.

Kolvin, I., Ounsted, C., Humphrey, M., & McNay, A. (1971). Studies in the childhood psychoses: II. The phenomenology of childhood psychoses. *British Journal of Psychiatry, 118*, 385-395.

Layne, C. (2007). Early identification of autism: Implications for counselors. *Journal of Counseling & Development, 85*, 110-114.

Levy, S. E., Mandell, D. S., Sands, J., Barbera, M., Brown, L., Brown, P., ... Wettlaufer, H. (2007). *Pennsylvania autism assessment and diagnosis expert work group: Supporting quality diagnostic practices for persons with suspected autism spectrum disorder.* Retrieved from Pennsylvania Department of Public Welfare website: http://www.dpw.state.pa.us

Lord, C., & Corsello, C. (2005). Diagnostic instruments in autistic spectrum disorders. In F. R. Volkmar, R. Paul, A. Klin, & D. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders* (3rd ed., pp. 730-771). Hoboken, NJ: Wiley.

Lord, C., Risi, S., Lanbrecht, L., Cook, E., Leventhal, B., DiLavore, ... Rutter, M. (2000). The Autism Diagnostic Observation Schedule- Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30*, 205-223.

Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (1999). *Autism Diagnostic Observation Schedule*. Los Angeles, CA: Western Psychological Services.

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. L. (2012). *Autism Diagnostic Observation Schedule, second edition*. Los Angeles, CA: Western Psychological Services.

Lotter, V. (1966). Epidemiology of autistic conditions in young children: I. Prevalence. *Social Psychiatry, 1*, 124-137.

Mawle, E., & Griffiths, P. (2005). Screening for autism in pre-school children in primary care: Systematic review of English language tools. *International Journal of Nursing Studies, 43*, 623–636. doi:10.1016/j.ijnurstu.2005.11.011

Mayo Clinic. (2012). *Child development chart: Preschool milestones*. Rochester, MN: Author. Retrieved from http://www.mayoclinic.com

Muller, E., & Markowitz, J. (2004). *Disability categories: State terminology, definitions, & eligibility determinations.* Alexandria, VA: National Association of State Directors of Special Education (NASDSE).

Retrieved from http://www.projectforum.org/docs/disability_categories.pdf

Myles, B., Jones-Bock, S., & Simpson, R. (2001). *Asperger Syndrome Diagnostic Scales*. Los Angeles, CA: Western Psychological Services.

National Research Council. (2001). *Educating children with autism.* Washington, DC: National Academy Press.

Noland, R., & Gabriels, R. (2004). Screening and identifying children with autism spectrum disorders in the public school system: The development of a model process. *Journal of Autism and Developmental Disorders, 34*, 265-277.

Norris, M., & Lecavalier, L. (2010). Screening accuracy of level 2 autism spectrum disorder rating scales. *Autism, 14*, 263-284.

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research, Methods, Instruments, & Computers, 32*, 396-402.

Oosterling, I., Roos, S., de Bildt, A., Rommelse, N., de Jonge, M., Visser, J., ... Buitelaar, J. (2010). Improved diagnostic validity of the ADOS revised algorithms: A replication study in an independent sample. *Journal of Autism and Developmental Disorders, 40*, 689-703. doi:10.1007/s10803-009-0915-0

Overton, T., Fielding, C., & Garcia de Alba, R. (2007). Differential diagnosis of Hispanic children referred for autism spectrum disorders: Complex issues. *Journal of Autism and Developmental Disorders, 37*, 1996-2007. doi:10.1007/s10803-006-0349-x

Palfrey, J., Singer, J., Walker, D., & Butler, J. (1987). Early identification of children's special needs: A study in five metropolitan communities. *Journal of Pediatrics, 111,* 651-659.

Pandey, J., Verbalis, A., Robins, D. L., Boorstein, H., Klin, A., Babitz, T., ... Fein, D. (2008). Screening for autism in older and younger toddlers with the Modified Checklist for Autism in Toddlers. *Autism, 12*, 513-535. doi:10.1177/1362361308094503

Pennsylvania Department of Education & Department of Public Welfare. (2009). *Pennsylvania learning standards of early childhood, revised second edition*. Retrieved from Office of Child Development and Early Learning website: http://www.pakeys.org/

Robertson, J. M., Tanguay, P. E., L'Ecuyer, S., Sims, A., & Waltrip, C. (1999). Domains of social communication handicap in autism spectrum disorder. *Journal of the*

American Academy of Child and Adolescent Psychiatry, 38, 738–745.

Robins, D. L. (2008). Screening for autism spectrum disorders in primary care settings. *Autism, 12*, 537-556. doi:10.1177/1362361308094502

Robins, D. L., & Dumont-Mathieu, T. M. (2006). Early screening for autism spectrum disorders: Update on the Modified Checklist for Autism in Toddlers and other measures. *Developmental and Behavioral Pediatrics, 27*, S111-S119.

Robins, D. L., Fein, D. F., & Barton, M. L. (1999). *The Modified Checklist for Autism in Toddlers*. Retrieved from http://www.mchatscreen.com

Robins, D. L., Fein, D. F., Barton, M. L., & Green, J. A. (2001). The Modified Checklist for Autism in Toddlers: An initial study investigating the early detection of autism and pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 31*, 131-144.

Robins, D. L., Pandey, J., Chlebowski, C., Carr, K., Zaj, J. L., Arroyo, M., ... Fein, D. A. (2010). *M-CHAT Best7: A new scoring algorithm improves positive predictive power of the M-CHAT*. Paper presented at the International Meeting for Autism Research, Philadelphia, PA.

Rutter, M. (1978). Diagnosis and definition. In M. Rutter & E. Schopler (Eds.), *Autism: A reappraisal of concepts and treatment* (pp. 1-25). New York, NY: Plenum Press.

Rutter, M. (2005). Genetic influences and autism. In F. R. Volkmar, R. Paul, A. Klin, & D. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders* (3rd ed., pp. 425–452). Hoboken, NJ: Wiley.

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

Rutter, M., Le Couteur, A., & Lord, C. (2003). *Autism Diagnostic Interview– Revised*. Los Angeles, CA: Western Psychological Services.

Schopler, E., Van Bourgondien, M. E., Wellman, G. J., & Love, S. R. (2010). *Childhood Autism Rating Scale, second edition.* Los Angeles, CA: Western Psychological Services.

Schwartz, I. S., & Davis, C. A. (2008). Best practices in effective services for young children with autistic spectrum disorders. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology V* (pp. 1517-1529). Bethesda, MD: National Association of School Psychologists.

Seigel, B. (2004). *Pervasive Developmental Disorders Screening Test-II (PDDST-II): Early childhood screeners for autistic spectrum disorders*. San Antonio, TX: Harcourt.

Shattuck, P. T., Durkin, M., Maenner, M. Newschaffer, C., Mandell, D. S., Wiggins, L., ... Cuniff, C. (2009). Timing of identification among children with an autism spectrum disorder: Findings from a population-based surveillance study. *Journal of the American Academy of Child and Adolescent Psychiatry, 48*, 474–483.

Shriver, M. D., Allen, K. D., & Matthews, J. R. (1999). Effective assessment of the shared and unique characteristics of children with autism. *School Psychology Review, 28*, 538-558.

Snow, A. V., & Lecavalier, L. (2008). Sensitivity and specificity of the Modified Checklist for Autism in Toddlers and the Social Communication Questionnaire in preschoolers suspected of having pervasive developmental disorders. *Autism, 12*, 627-644. doi:10.1177/1362361308097116

South, M., Williams, B., McMahon, W., Owley, T., Filipek, P., Shernoff, E., ... Ozonoff, S. (2002). Utility of the Gilliam Autism Rating Scale in research and clinical populations. *Journal of Autism and Developmental Disorders, 32*, 593-599.

Stone, W. L., Coonrod, E. E., & Ousley, O. Y. (2000). Brief report: Screening Tool for Autism in Two-Year-Olds (STAT): Development and preliminary data. *Journal of Autism and Developmental Disorders, 30*, 607–612.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Pearson.

Waterhouse, L., Wing, L., Spitzer, R., & Siegel, B. (1992). Pervasive developmental disorders: From DSM-III to DSM-III-R. *Journal of Autism and Developmental Disorders, 22*, 525-549.

Watkins, M. W. (2000). Monte Carlo PCA for Parallel Analysis [Computer Software]. University Park, PA: Marley W. Watkins.

Wetherby, A. M., Allen, L., Cleary, J., Kublin, K., & Goldstein, H. (2002). Validity and reliability of the Communication and Symbolic Behavior Scales Developmental Profile with very young children. *Journal of Speech Language and Hearing Research, 45*, 1202–1218.

Wetherby, A. M., & Woods, J. (2004). *Systematic observation of red flags for autism spectrum disorders in young children.* Tallahassee: Florida State University Press.

Wilkinson, L. A. (2010). *A best practice guide to assessment and intervention for autism and Asperger syndrome in schools.* London, UK: Jessica Kingsley.

Wilkinson, L. A. (2011, October). Identifying students with autism spectrum disorders: A review of selected screening tools. *NASP Communiqué, 40*(2), 1, 31-33.

Wing, L. (1981). Asperger's syndrome: A clinical account. *Psychological Medicine, 11*, 115-129.

Wing, L. (1997). The history of ideas on autism: Legends, myths, and reality. *Autism, 1*(1), 13-23.

Wing, L. (2005). Problems of categorical classification systems. In F. R. Volkmar, R. Paul, A. Klin, & D. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders* (3rd ed., pp. 583–605). Hoboken, NJ: Wiley.

Wing, L., & Potter, D. (2009). The epidemiology of autism spectrum disorders: Is the prevalence rising? In S. Goldstein, J. A. Naglieri, & S. Ozonoff (Eds.), *Assessment of autism spectrum disorders* (pp. 18-54). New York, NY: Guilford.

Yama, B., Freeman, T., Graves, E., Yuan, S., & Campbell, K. (2012). Examination of the properties of the Modified Checklist for Autism in Toddlers (M-CHAT) in a population sample. *Journal of Autism and Developmental Disorders, 42*, 23-34. doi:10.1007/s10803-011-1211-3

Yeargin-Allsop, M., Rice, C., Karapurkar, T., Doernberg, T., Boyle, C., & Murphy, C. (2003). Prevalence of autism in a US metropolitan area. *Journal of the American Medical Association, 289*, 49-55.

Yell, M. L., & Drasgow, E. (2000). Litigating a free and appropriate public education: The Lovaas hearings and cases. *Journal of Special Education, 33*, 205–215.

M-CHAT

Please fill out the following about how your child usually is. Please try to answer every question. If the behavior is rare (e.g., you've seen it once or twice), please answer as if the child does not do it.

1. Does your child enjoy being swung, bounced on your knee, etc.?	Yes	No
2. Does your child take an interest in other children?	Yes	No
3. Does your child like climbing on things, such as up stairs?	Yes	No
4. Does your child enjoy playing peek-a-boo/hide-and-seek?	Yes	No
5. Does your child ever pretend, for example, to talk on the phone or take care of dolls, or pretend other things?	Yes	No
6. Does your child ever use his/her index finger to point, to ask for something?	Yes	No
7. Does your child ever use his/her index finger to point, to indicate interest in something?	Yes	No
8. Can your child play properly with small toys (e.g. cars or bricks) without just mouthing, fiddling, or dropping them?	Yes	No
9. Does your child ever bring objects over to you (parent) to show you something?	Yes	No
10. Does your child look you in the eye for more than a second or two?	Yes	No
11. Does your child ever seem oversensitive to noise? (e.g., plugging ears)	Yes	No
12. Does your child smile in response to your face or your smile?	Yes	No
13. Does your child imitate you? (e.g., you make a face-will your child imitate it?)	Yes	No
14. Does your child respond to his/her name when you call?	Yes	No
15. If you point at a toy across the room, does your child look at it?	Yes	No
16. Does your child walk?	Yes	No
17. Does your child look at things you are looking at?	Yes	No
18. Does your child make unusual finger movements near his/her face?	Yes	No
19. Does your child try to attract your attention to his/her own activity?	Yes	No
20. Have you ever wondered if your child is deaf?	Yes	No
21. Does your child understand what people say?	Yes	No
22. Does your child sometimes stare at nothing or wander with no purpose?	Yes	No
23. Does your child look at your face to check your reaction when faced with something unfamiliar?	Yes	No

M-CHAT Scoring Instructions

A child fails the checklist when 2 or more critical items are failed OR when any three items are failed. Yes/no answers convert to pass/fail responses. Below are listed the failed responses for each item on the M-CHAT. Bold capitalized items are CRITICAL items.

Not all children who fail the checklist will meet criteria for a diagnosis on the autism spectrum. However, children who fail the checklist should be evaluated in more depth by the physician or referred for a developmental evaluation with a specialist.

1. No	6. No	11. Yes	16. No	21. No
2. NO	**7. NO**	12. No	17. No	22. Yes
3. No	8. No	**13. NO**	18. Yes	23. No
4. No	**9. NO**	**14. NO**	19. No	
5. No	10. No	**15. NO**	20. Yes	

M-CHAT Best7 Scoring Instructions

A child screens positive, or shows **Risk for Autism**, when 2 or more "Best7" items are failed OR when any three items are failed. If fewer than 2 "Best7" items are failed, and fewer than 3 total items are failed, the result is **Low Risk for Autism**. The design of M-CHAT Best7 is to retain high sensitivity with a low false-positive rate for Autism concern. If the result of the checklist is "Risk for Autism" the corresponding M-CHAT Follow-up Interview™ should be given to obtain the most accurate responses.

Yes/no answers convert to pass/fail responses. Below are listed the failed responses for each item on the M-CHAT. **BOLD CAPITALIZED** items are "Best7" items.

Not all children who fail the checklist will meet criteria for a diagnosis on the autism spectrum. However, children who screen positive on the M-CHAT should be evaluated in more depth by the physician or referred for a developmental evaluation with a specialist.

1. No	6. No	11. Yes	16. No	21. No
2. NO	**7. NO**	12. No	17. No	22. Yes
3. No	8. No	13. No	18. Yes	23. No
4. No	**9. NO**	**14. NO**	19. No	
5. NO	10. No	**15. NO**	**20. YES**	

Appendix C, Permission to Duplicate the M-CHAT

May 6, 2011

Department of Psychology
Georgia State University
Atlanta, GA 30302

Dear Dr. Robins,

I am completing a doctoral dissertation at the Indiana University of Pennsylvania. The dissertation will be microfilmed by UMI Dissertation Publishing (ProQuest Information and Learning) and copies will be sold on demand. I would like permission to duplicate the following for research use:

Title: The Modified Checklist for Autism in Toddlers (M-CHAT)

Copyright: Robins, Fein, & Barton, 1999

Authors: Diana Robins, Deborah Fein, & Marianne Barton

Material to be Duplicated: All questions (1-23) from the M-CHAT

Use: To be included in my dissertation entitled "Differences in performance on the *Modified Checklist for Autism in Toddlers* between students with Autism and other developmental delays in a preschool-age special education sample".

You can mail the permission to JoAnna Cogan-Ferchalk, 444 Hill Road, Wernersville, PA, 19565, or email a scanned copy to j.r.cogan@iup.edu

Please let me know what conditions, if any, apply to this use.

Sincerely,

JoAnna R. Cogan-Ferchalk
Doctoral Candidate
Department of Educational and School Psychology

Permission granted _ 5/9/2011

____ ______________

Signature Date

Conditions, if any: M-CHAT must be reprinted in its entirety, without altering title, instructions, item content, item order, and copyright notice. If a website is referenced, it should be www.mchatscreen.com.

Appendix D, Permission to Access Archival Data

helping schools excel | helping children learn

BCIU
Berks County Intermediate Unit
an educational service agency

April 5, 2011

Berks County Intermediate Unit
1111 Commons Blvd PO Box 16050
Reading, PA 19612

To Whom It May Concern:

I have read the research proposal submitted by JoAnna Cogan-Ferchalk and understand that she intends to study the Modified Checklist for Autism in Toddlers (M-CHAT) for preschool-age students enrolled in the BCIU Early Intervention Program. I hereby agree to allow the proposed research study to take place in the Berks County Intermediate Unit and I further allow Ms. Cogan-Ferchalk access to the archival data needed to complete this study. I understand that confidentiality of this data will be maintained by disassociating participants' identifying information from all data collected. All data will be analyzed in group format and no student names will be attached to individual scores.

If you have any questions or concerns, you can contact me at 610/987-8511. Thank you.

Sincerely,

Jill M. Hack

Jill M. Hackman, Ed.D.
Director
Office of Early Childhood and Student Services

JMH/jfw

1111 Commons Boulevard
P.O. Box 16050
Reading PA 19612-6050
610-987-BCIU (2248)
FAX 610-987-8400 | TTY 610-987-8401
www.berksiu.org