

2004

Reliability and concordance of the Childhood Autism Rating Scale and DSM-IV in adults with severe and profound mental retardation

Noha Farrah Minshawi

Louisiana State University and Agricultural and Mechanical College

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_theses



Part of the [Psychology Commons](#)

Recommended Citation

Minshawi, Noha Farrah, "Reliability and concordance of the Childhood Autism Rating Scale and DSM-IV in adults with severe and profound mental retardation" (2004). *LSU Master's Theses*. 1609.
https://digitalcommons.lsu.edu/gradschool_theses/1609

This Thesis is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Master's Theses by an authorized graduate school editor of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.

RELIABILITY AND CONCORDANCE
OF THE CHILDHOOD AUTISM RATING SCALE AND DSM-IV
IN ADULTS WITH SEVERE AND PROFOUND MENTAL RETARDATION

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Master of Arts

in

The Department of Psychology

By
Noha Farrah Minshawi
B.S., The Ohio State University, 2001
May, 2004

Table of Contents

Abstract.....	iii
Introduction.....	1
Dual Diagnosis.....	7
Autism	10
Behavioral Assessment of Autism	22
Childhood Autism Rating Scale (CARS)	25
Interrater Agreement and Staff Accuracy.....	28
Autism and Mental Retardation	31
Purpose.....	36
Method.....	39
Results.....	43
Discussion.....	52
References.....	61
Appendix A: DSM-IV Checklist.....	72
Appendix B: CARS and DSM-IV Checklist Administration Instructions.....	73
Vita.....	75

Abstract

Researchers and clinicians have long treated autism as though it were a disorder that only affected children. As a result, little literature is available on the diagnosis, assessment, and treatment of adults with autism. One of the first issues to address in this population is diagnosis. While several rating scales and diagnostic systems exist for surveying autistic behavior in children, researchers have not demonstrated the reliability of these scales for adults. The present study focused on two commonly used instruments, the Childhood Autism Rating Scale (CARS) and the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV). In the first objective, reliability of the two instruments was established using two types of informants: direct-care staff and trained mental health professionals. Test-retest reliability of the CARS using direct-care staff informants was good and better than the test-retest reliability of the DSM-IV criteria, which was acceptable. Interrater reliability between two direct-staff, as well as between direct-care staff and trained mental health professionals was unacceptable for clinical use. The second objective of the study was to assess concordance between CARS and DSM-IV diagnoses. For both direct-care staff and trained mental health professionals, CARS and DSM-IV diagnoses concurred at a rate greater than that expected by chance. In the final portion of the study, CARS scores and DSM-IV diagnoses for individuals with profound mental retardation (PMR; n = 46) and severe mental retardation (SMR; n = 46) were compared as a preliminary step towards determining the appropriateness of these two instruments in individuals with mental retardation. The PMR group had significantly higher CARS scores and significantly more DSM-IV diagnoses than the SMR group. Results and implications of the study are discussed.

Introduction

Mental retardation has existed throughout history. In the middle ages, physical disorders were given more attention than mental conditions. At this time, doctors began reporting cases of hydrocephalus and epilepsy (Scheerenberger, 1983). In the sixteenth century, Paracelsus, a Swiss physician, made the first distinction between mental illness and mental retardation (Fiedler, 1978). Paracelsus noted that a high degree of variability existed in each. Declared to be “idiots” or “feeble-minded,” individuals with mental retardation began receiving more attention during this time, although little treatment occurred; persons with mental retardation were deemed untrainable.

The 1800s proved to be a period of great optimism and progress in the field of mental retardation (Scheerenberger, 1983). A physician named W.J. Little was the first to draw attention to the deleterious effects of premature birth (Leakey & Lewin, 1977). Little provided case studies that demonstrated the consequences of labor difficulties, hypoxia, and mechanical injuries during birth. Around the same time, Edouard Seguin divided idiocy (mental retardation) into four categories: idiocy, imbecility, backwardness or feeble-mindedness, and simpleness (Abt, 1965). Seguin was also the first advocate for the training of these individuals, suggesting work that required simple, repetitive movements.

At the end of the 19th and beginning of the 20th centuries interest in mental retardation skyrocketed. The establishment of professional organizations, journals, and international conferences marked this period. In 1877, the American Association of Mental Deficiency (AAMD) officially endorsed its first definition of mental retardation (Scheerenberger, 1983). The AAMD stated that idiocy was a lack of natural development of the mental and moral (social) powers, usually accompanied by physical defect. The other hallmark of this time was the advent

of special education in the public schools. The Providence, Rhode Island public school system is often given credit for establishing the first public special education classroom (Elkin, 1977).

Intended for the mildly mentally retarded, special education services included both regular academic courses and physical education (Scheerenberger, 1983).

The creation of public special education increased the need for standardization in the psychological measurement of mental retardation. In 1905, the first standardized test of intelligence, the Binet-Simon Individual Tests of Intelligence, was created (Scheerenberger, 1983). The Binet-Simon was intended to distinguish between normal and below normal school-aged children. Scoring of the test fell into three levels of mental retardation: idiocy, imbecility, and moronity. In addition to test standardization, the beginning of the 20th century was characterized by an increase in social control over individuals with mental retardation. Marriage laws prohibited the marriage of two people with mental retardation, and many people endorsed the use of sterilization of persons with mental retardation along with the insane, criminals, and sex offenders. These negative attitudes carried over into the proliferation of institutions for the mentally retarded.

The next period of reform in the history of mental retardation came in the 1950s and 1960s. The revised Stanford-Binet Intelligence Test provided new terminology for the classification of mental retardation. The new classification consisted of borderline (IQ of 83-67), mild (IQ of 66-50), moderate (IQ of 49-33), severe (IQ of 32-16), and profound (IQ of 16) mental retardation. Researchers focused on behavioral aspects of mental retardation, such as learning, motivation, and attention, and with this research came a rapid expansion of special education classrooms across the country.

In 1961 the AAMD revised its definition of mental retardation to the "...sub average

general intellectual functioning, which originates in the developmental period and is associated with impairment in adaptive behavior" (Hayes, 1969, p.17). Borderline mental retardation was excluded from this definition, making mild mental retardation an IQ of 67-52, moderate an IQ of 51-36, severe an IQ of 35-20, and profound an IQ of 19 and below. In addition, this definition contributed a new feature to the understanding of mental retardation: impairment in adaptive behavior. Adaptive behavior consists of basic motor and self-help skills, learning, and social adjustment. Researchers began focusing on ways to measure adaptive behavior, and eventually two rating scales were introduced: the AAMD Adaptive Behavior Scale (ABS) (Nihira, Foster, Shellhaas, & Leland, 1969) and the Vineland Adaptive Behavior Scale (VABS) (Sparrow, Balla, & Cicchetti, 1984).

Current Diagnosis of Mental Retardation

The history of the classification of mental retardation has lead to our current understanding of the disorder. In 1994, the *Diagnostic and Statistical Manual of Mental Disorders Fourth Edition* (DSM-IV; APA, 1994) defined mental retardation as significantly sub average intellectual functioning accompanied by limitations in at least two of eight domains: communication, self-care, home living, social/interpersonal skills, work, leisure, health, and safety. Onset of these deficits must occur prior to the age of 18. Sub average intelligence is characterized by an IQ of approximately 70 or below on a standardized intelligence test, and adaptive behavior is defined as a person's effectiveness in meeting the standards expected for his or her age.

Mental retardation is currently divided into the same four categories as proposed by the AAMD in 1961. An individual meets the diagnosis of mild mental retardation if (s) he has an IQ of approximately 70 to 50-55. Persons with mild mental retardation make up 85% of those with

the disorder (APA, 1994). Children with mild mental retardation usually develop appropriate social and communicative skills in preschool and are not identifiable until their late teenage years. As adults, these individuals often acquire sufficient social and vocational skills as to require minimal assistance to live independently or in a group-home setting (APA, 1994).

The next category is moderate mental retardation. Moderate mental retardation is diagnosed when IQ falls between 55-50 and 35-40. Persons with moderate mental retardation make up 10% the disorder's population (APA, 1994). An individual with moderate mental retardation rarely progresses past the second-grade level of schooling, and often exhibits difficulty in social situations. These individuals can benefit greatly from social and vocational training.

An IQ of 35-40 to 20-25 characterizes the third category, severe mental retardation. Severe mental retardation constitutes only 2-3% of individuals with mental retardation. A child with severe mental retardation will usually acquire little or no communicative speech, but may eventually learn limited speech and basic self-help skills.

The final category is profound mental retardation. Profound mental retardation constitutes the remaining 1-2% of the population. Profound mental retardation is diagnosed when IQ falls below 20 or 25, and is usually accompanied by deficits in sensorimotor functioning, communication, and motor development (APA, 1994). For many persons with profound mental retardation, a specific neurological condition can be identified that accounts for the mental retardation. Most of these individuals require highly supervised and individualized care (APA, 1994).

Prevalence and Etiology of Mental Retardation

Mental retardation is believed to occur at a rate of approximately 1% in the general

population (APA, 1994). The 1994/1995 National Health Interview Survey (Larson et al., 2001) estimated the rate of mental retardation to be .78%, the rate of developmental disabilities to be 1.13%, and the rate of combined mental retardation and developmental disabilities to be 1.49%. Bernsen (1976) and Darragh (1982) reported the prevalence of an IQ below 50 to be three to four per 1,000. The rates vary when this category is divided into moderate, severe, and profound levels of mental retardation. The prevalence of moderate impairment has been found to be two per 1,000, and that of severe impairment to be 1.3 per 1,000 (Fishbach & Hull, 1982; McQueen, Spence, Garner, Pereira, & Winsor, 1987). Reported prevalence of profound mental retardation is 0.4 per 1,000 (Baird & Sadovnick, 1985; McQueen et al., 1987).

The prevalence of mental retardation across gender has also been investigated. The DSM-IV reports a male to female ratio of 1.5:1 (APA, 1994), which is similar to that reported by other studies (e.g. Laxova, Ridler, & Bowen-Bravery, 1977; Lindsey & Russell, 1981). The reasons for an increased rate of mental retardation among males are not clear. Further study into the role of gender, degree of retardation, and etiology are necessary to begin to understand the relationship between gender and mental retardation (McLaren & Bryson, 1987).

The final subject under review in this section is the etiology of mental retardation, which is as broad as the disorder itself. In approximately 30-40% of individuals with mental retardation there is no known etiology (APA, 1994). For the remaining portion of the population, etiological factors include heredity, genetics, prenatal and perinatal conditions, and general medical conditions acquired in infancy and early childhood.

Several genetic causes of mental retardation exist. Some heritable forms of mental retardation include Tay-Sachs disease, tuberous sclerosis, fragile X syndrome, and phenylketonuria. In addition, early alterations in embryonic development, such as those seen in

Down's syndrome and Fetal Alcohol Syndrome, can also predispose an individual to mental retardation. Other common predisposing factors are conditions that arise during prenatal and perinatal development. These conditions include fetal malnutrition, prematurity, hypoxia, trauma, and viral infection. Prenatal factors are thought to be the cause of 20% to 30% of the cases of mental retardation (Gustavson, Hagberg, Hagberg, & Sars, 1977; McQueen et al., 1986), and are considered to be more common than perinatal and postnatal factors combined (McLaren & Bryson, 1987). Finally, general medical conditions that occur in infancy and early childhood, such as infection, trauma, and lead poisoning, are also contributors.

Knowledge regarding the etiology of mental retardation is far from complete. The issue is further clouded by the fact that a large portion of the population, as much as 50%, present with more than one potential causal factor (Gustavson et al., 1977; McQueen et al., 1986). Researchers continue to focus on the biological factors associated with mental retardation, along with psychosocial variables. In addition to these numerous causes of mental retardation, these persons are also at increased risk for mental health conditions. This phenomenon, referred to as dual diagnosis, will be reviewed next.

Dual Diagnosis

The next topic under consideration is dual diagnosis of persons with mental retardation.

A wide variety of mental illnesses have been reported in individuals with mental retardation (Matson, 1985), including emotional disorders (Menolascino, 1977), depression (Menolascino, 1988), anxiety (Borthwick-Duffy, 1994), and schizophrenia (Menolascino, 1988; Heaton-Ward, 1977). Psychopathology in individuals with mental retardation presents an increased challenge to clinicians and is frequently the cause of institutionalization (Hill & Bruninks, 1984).

Rutter and colleagues (1970) investigated the prevalence of psychopathology in mental retardation in a series of studies on childhood psychopathology on the Isle of Wight in England (Rutter, 1989; Rutter, Graham, & Yule, 1970; Rutter, Tizard, & Whitmore, 1970). Parent and teacher report, as well as psychiatric measures, were collected on cohorts of 9, 10, and 11-year-old children. The overall prevalence of psychopathology in the population was 7%, but for those children with an IQ below 70, the prevalence rate jumped to 30-42%, an increase of nearly five times that of the general population. Similarly, Matson and Frame (1986) estimated the prevalence of psychopathology in children with mental retardation to be two to three times greater than that of children without mental retardation. Another large-scale study of 1,507 residents in an institution for the mentally retarded found that serious psychiatric conditions occurred in 7.3% of the sample (Wright, 1982).

Within mental retardation, the prevalence of psychopathology varies. Evidence points towards a relationship between level of mental retardation and occurrence of psychiatric disorders. Individuals with mild mental retardation evince the greatest rates of psychopathology (Borthwick-Duffy & Eyman, 1990). The cause for this relationship is not well understood.

However, these results may not demonstrate a problem inherent in individuals with mild mental retardation, but instead a problem in current diagnostic and classification criteria.

Assessment

In order to address these diagnostic issues, two checklists have been developed specifically for use with individuals with mental retardation. The *Diagnostic Assessment for the Severely Handicapped-II* (DASH-II), and *Assessment for Dual Diagnosis* (ADD) are two commonly used instruments for assessing psychopathology in persons with mental retardation. The DASH-II was designed to assess behavioral and psychiatric symptoms in individuals with severe and profound mental retardation (Matson, 1995b). The 84 items of the DASH-II represent the following 13 diagnostic categories: 1) Anxiety; 2) Depression; 3) Mania; 4) PDD/Autism; 5) Schizophrenia; 6) Stereotypies; 7) Self-Injurious Behavior; 8) Elimination Disorders; 9) Eating Disorders; 10) Sleep Disorders; 11) Sexual Disorders; 12) Organic Syndromes; 13) Impulse Control. Finally, the ADD is an instrument that provides diagnostic information on psychopathology in persons with mild or moderate mental retardation (Matson & Bamburg, 1998). Items on the ADD were derived from the DSM-IV and other research on psychopathology in mental retardation. The diagnostic categories represented on the ADD are: 1) Mania; 2) Depression; 3) Anxiety; 4) Post-Traumatic Stress Disorder; 5) Substance Abuse; 6) Somatoform Disorders; 7) Dementia; 8) Conduct Disorder; 9) Pervasive Developmental Disorder; 10) Schizophrenia; 11) Personality Disorders; 12) Eating Disorders; 13) Sexual Disorders. These two assessment tools have provided clinicians and researchers with valuable insight into the diagnosis of psychopathology, especially in those individuals with more severe forms of mental retardation. However, they are used as initial methods to narrow the focus to a

specific set of problems or a disorder. Scales for specific disorders, such as autism, are needed as a step towards accurate diagnoses and targets for intervention.

Autism

The disorder under examination in the present study is autism. In this section the history of autism will be discussed, as well as core features, prevalence, etiology, and adult outcome of the disorder.

Description

In 1943 Leo Kanner offered the first description of what he called infantile autism. Kanner observed in 11 children a similar pattern of behaviors, which included abnormal language development and use, social skills deficits, and insistence on sameness.

Initial reports of language by the parents of the 11 children were similar. The children developed language according to normal milestones, and began memorizing and repeating nursery rhymes, poems, and songs at a very young age. However, as the children aged they failed to begin asking or answering questions. Kanner described the language of these children as being used for a function other than communication. The children's language was characterized as literal and inflexible, and many persons with autism were unable to generalize and transfer an expression from one object or situation to another. Language was also irrelevant at times, with the children repeating phrases inappropriately that they had previously heard. Personal pronouns were also used incorrectly, typically being used when "quoting," (i.e. echoing,) something previously heard. In addition, the children also produced nonfunctional sounds in a very repetitive manner. A few of the children failed to acquire communicative speech, Kanner referred to these children as mute (Kanner, 1943).

Reports of social skills difficulties were spread throughout Kanner's original accounts of autism. Parent reports of infancy and early childhood contained abundant references to their children's lack of interest in the comings and goings of the people in their homes and the other

children in their neighborhoods. Parents described the children as being self-sufficient, oblivious, hypnotized, and happiest when left alone. Kanner noted that upon entering a room the children paid no attention to the people in the room and instead went straight to the objects. When forced to interact with other people, these children showed annoyance, resentment, and often times anger. The little interaction that could be elicited from the children was devoid of eye contact and reciprocity.

Finally, Kanner made detailed mention of the children's insistence on sameness. Most of the children exhibited a marked limitation in spontaneous activity. Toys were played with the same way each time, blocks were arranged precisely by color or size, and beads were strung in exactly the same order. When these rituals were interrupted, or routines were changed, some of the children became uncomfortable or angry (Kanner, 1943). Several of the children became disturbed at the sight of something incomplete or broken. In addition, Kanner noted that several of the children possessed irrational fears. Riding a tricycle horrified one child; another child was frightened of mechanical objects (including his mother's egg beater).

Kanner viewed the combination of language and social deficits, coupled with an insistence on sameness, stereotyped patterns of behavior, echolalia, and obsessive behavior, as a disorder that could be differentiated from other childhood disorders. Moreover, Kanner did not incorporate intellectual deficiency into his description of autism because he did not believe these children were mentally retarded. Kanner viewed the children's good rote memory, intelligent parents, and lack of physical deformity as evidence that autistic children were of average, if not above average, intelligence.

Several years after Kanner's description of infantile autism, another researcher, Creak, developed a set of criterion for the identification of early childhood psychosis (Creak, 1961).

Creak's nine characteristics were: 1) gross impairment of emotional relationships; 2) age inappropriate lack of awareness of personal identity; 3) pathological preoccupation with certain objects or their characteristics, without regard for function; 4) resistance to environmental change and effort to maintain or restore sameness; 5) abnormal perceptual experience; 6) anxiety; 7) loss of speech or failure to acquire or develop language; 8) distorted pattern of motility; 9) history of serious retardation. While many of his characteristics overlapped with Kanner's, Creak, like other researchers at the time (i.e. Fish, 1976), believed that these behaviors were actually a form of childhood schizophrenia. Despite this, Creak's nine features were incorporated into many subsequent descriptions of autism and commonly used autism assessment instruments (e.g. Krug, Arick, & Almond, 1979; Schopler, Reichler, DeVellis, & Daly, 1980).

In 1978, Rutter attempted to clarify the definition of autism. Rutter believed that autism was a distinct syndrome because the behaviors observed occurred with uniformity across all subjects and were specific to autism. Thus, autism could be differentiated from other developmental disorders. According to Rutter, three broad categories of symptoms could be found in autism. These categories of behaviors were: impaired social relations, delayed and/or abnormal language development, and compulsive behavior or insistence on sameness (Rutter, 1978). This triad of behaviors was influential in the development of autism as an accepted clinical diagnosis.

With Rutter's synthesis of Kanner's work, autism, (called infantile autism,) was included in the Diagnostic and Statistical Manual of Mental Disorders, Third Edition (DSM-III; APA, 1980). In DSM-III, autism was included among a new class of disorders named pervasive developmental disorders (PDD). Parents were concerned that their children not be classified as

having a traditional thought disorder, such as schizophrenia. Thus, the term PDD was used. In order for a child to meet a diagnosis of infantile autism (s) he had to exhibit a pervasive lack of social relationships and language deficits, before the age of 30 months. In line with Kanner's belief that autism affected only young children, the DSM-III included a diagnosis of "residual" autism for those individuals who had met the criteria for infantile autism at one point, but no longer exhibited the symptoms. The official recognition of autism provided vast research opportunities, but the DSM-III was wrought with problems, including an overemphasis on the "infantile" aspect of the disorder (Volkmar, Bregman, Cohen, & Cicchetti, 1988), overly stringent criteria, and a lack of consideration of issues associated with developmental change (Volkmar, Cicchetti, Bregman, & Cohen, 1992a). These issues were addressed in the subsequent revision of the DSM-III (The Diagnostic and Statistical Manual of Mental Disorders, Third Edition, Revised; DSM-III-R; APA, 1987).

The DSM-III-R replaced infantile autism with autism, recognizing that individuals with autism continued to exhibit symptomology beyond childhood. The category of "residual" autism was also removed, allowing the disorder to be diagnosed in individuals of any age or developmental level. In addition, the criteria for autism were broadened to encompass potential developmental changes. With this expansion in criteria came a high rate of false positive identifications (Factor, Freeman, & Kardish, 1989; Hertzig, Snow, New, & Shapiro, 1990; Volkmar, Cicchetti, Bregman, & Cohen, 1992b). Clarification of these false positives, as well as the reasonable range of symptom inclusion and rationale for the use of historical information, provided the basis for the field trials for DSM-IV (Volkmar et al., 1994).

The DSM-IV field trials compared the DSM-III-R criteria with that of the International Classification of Diseases, 10th Edition (ICD-10; WHO, 1996). The ICD-10 diagnosis of autism

is made on the basis of delayed or deviant communication, social interaction and/or play, qualitative impairment in social interaction, and restricted and stereotyped patterns of behavior or interests. Symptom onset must be prior to age three. Subjects in the field trial were 454 individuals diagnosed with autism by an experienced clinician, 240 diagnosed with other pervasive developmental disorders, and 283 with conditions other than pervasive developmental disorders. Items from the ICD-10, DSM-III and DSM-III-R were compiled into a random checklist and completed by an independent rater for each subject. Clinical diagnoses were determined using the scoring rules for each individual rating system. Sensitivity, specificity, positive, negative and total predictive values, and agreement with clinician (kappa coefficient) were calculated for all three scales. Volkmar et al. (1994) suggested that the ICD-10 was superior to both the DSM-III and DSM-III-R because it offered a better combination of sensitivity and specificity over the age ranges and developmental levels. The criticism of the ICD-10 was that it was too detailed, both in number of items and length of criteria. Using information from the reliability of individual items, four of the original ICD-10 criteria were eliminated because they had a low frequency of endorsement or were duplicate information from other criteria. The field trial resulted in our current DSM-IV definition of autism.

The DSM-IV's (APA, 1994) current definition of autism states that the essential features are impaired development of social interaction and communication, and restricted range of interest and activities. The current diagnosis of autism revolves around these three criterion while providing more detailed definitions. The first criterion of the diagnosis of autism is social skills deficits. An individual must exhibit impaired social interaction manifested by at least two of the following behaviors: marked impairment in the use of several nonverbal forms of communication, (such as eye contact, facial expression and gaze,) failure to establish

developmentally appropriate peer relations, lack of spontaneous seeking of shared interests with others, or lack of social or emotional reciprocity. The second criterion, qualitative impairment in communication, is endorsed if the individual manifests at least one of the following behaviors: lack of or delay in the development of speech, inability or impairment in initiating or sustaining conversation, stereotyped or repetitive use of language, or lack of imaginative or imitative play. The third criterion, restricted, repetitive, and stereotyped patterns of behavior is endorsed when at least one of the following behaviors is exhibited: preoccupation with one or more stereotyped patterns of interest, inflexible adherence to specific, nonfunctional routines, stereotyped and repetitive motor behaviors, or preoccupation with parts of objects. In addition to these criterions, the onset of abnormal function must be prior to age three. Finally, the DSM-IV cautions that the clinician or researcher should determine whether the symptoms are better accounted for by either Rhett's Disorder or Childhood Disintegrative Disorder.

Core Features of Autism

To summarize current knowledge of the behavioral characteristics of autism, research has identified several core features of the disorder. The core features of autism are social skills deficits, language deficits, insistence on sameness, and abnormal response to sensory stimuli (Sturmey & Sevin, 1994).

Social skills deficits. The first core feature of autism is qualitative impairment of social interaction and relationships. Researchers have supported social skills deficits as constituting the central feature of autism (Hobson, 1989; Mundy & Sigman, 1989; Ungerer, 1989; Volkmar, 1987). In infants, this may be exhibited as rigidity, failure to seek physical comfort from other people (DeMeyer, 1979), and failure to develop normal attachment to parents and caregivers (Sturmey & Sevin, 1994). These children may appear to be just as comfortable with strangers as

they are with their own parents. In addition, autistic children fail to develop reciprocal eye contact and social smiling (Volkmar, 1987).

In a retrospective study of parents of children with autism, 85% of parents reported that their children ignored people, 90% seemed hard to reach, and 76% avoided eye contact (Ornitz, Guthrie, & Farley, 1978). As children with autism develop, social skills deficits remain (Rutter & Garmezy, 1983). School-aged autistic children often lack specific peer friendships and rarely engage in peer play (Howlin & Rutter, 1987). Normal displays of affection or empathy are also uncommon (Sturmey & Sevin, 1994). This impairment in social skills has been used to differentiate children with autism from children with mental retardation. Rodrigue, Morgan, and Geffken (1991) found that children with autism exhibited significantly greater socialization deficits than children with mental retardation.

Comparatively little is known about the social skills of adults with autism (Njardvik, Matson, & Cherry, 1999). One study reported that adults with autism between the ages of 21 and 35 displayed fewer adaptive skills than persons with mental retardation in the same age range (Jacobson & Ackerman, 1990). Another study compared the social skills of adults with mental retardation and comorbid autism to adults with mental retardation only(Njardvik et al., 1999). Social skills were measured by the Matson Evaluation of Social Skills in Individuals with sEvere Retardation (MESSIER) (Matson, 1995a). Adults with autism and mental retardation combined scored significantly lower than those with only mental retardation on the general positive and positive nonverbal subscales of the MESSIER.

Language deficits. Poor communicative skills are a hallmark of autism (Rutter, 1978). In fact, many children with autism never acquire functional language skills. When speech does develop, it is often marked with irrelevant content and stereotyped and repetitive vocalizations.

Kanner (1943) noted several communication abnormalities in the children he observed, including mutism, an emphasis on literal meaning, echolalia, improper use of pronouns, and use of language for purposes other than communication. Studies on the development of nonverbal communication in autism have demonstrated a lack of preverbal pointing (Bartak, Rutter, & Cox, 1975) and failure to use “showing” gestures (Curcio, 1978). However, inappropriate use of language and inability to use language for social communication are more characteristic of autistic language deficits (Howlin & Rutter, 1987). Finally, the ability to sustain conversation and produce spontaneous language is greatly limited in persons with autism (Matson, Sevin, Fridley, & Love, 1990; Stone & Caro-Martinez, 1990).

Insistence on sameness. Kanner (1943) observed that many of the children he studied showed distress over changes in their environment or routine. As a result, many aspects of the lives of children with autism become ritualized (Rutter, 1978). This insistence on sameness can be manifested as an obsession with particular items or places, as well as a highly restricted range of interest. The play of children with autism is characteristically rigid and lacking in variety and imagination (Rutter, 1978). Changes in environment, such as rearranging furniture, can result in extreme emotional reactions (Kanner, 1951). In addition, ritualistic and compulsive behaviors are common in autism and may be expressed as touching compulsions or number rituals, among others (Rutter, 1978).

Abnormal response to sensory stimuli. The final core feature of autism is abnormal response to sensory stimuli. Far less research is available regarding this feature, but both hypersensitivity and hyposensitivity have been reported (Brooker & Mareth, 1982; Ritvo & Freeman, 1977; Schopler, Reichler, & Renner, 1988). Clinical accounts of children with autism responding with undue distress to low volume sounds or being unable to differentiate between

verbal communication and other environmental noises are common. Insensitivity to pain has also been noted in several descriptions of autism (Mahler, 1952; Rimland, 1964)

Prevalence of Autism

With the features of autism firmly established in the literature, attention has shifted to the prevalence of the disorder. The most comprehensive prevalence study to date was conducted by Lotter in the late 1960s (1966; 1967). Lotter surveyed all 8 to 10-year-old children in a county near London, England. This survey presented a prevalence rate of autism of 4.5 per 10,000. The most extensive prevalence studies in the United States have resulted in similar rates of 3.1 (Treffert, 1970) and 4.0 (Ritvo, Freeman, Mason-Brothers, Mo, & Ritvo, 1985) cases per 10,000. The DSM-IV (APA, 1994) stated that the median reported rate in epidemiological studies is 5 cases per 10,000, with rates ranging from 2 to 20 per 10,000.

Researchers of all the epidemiological studies to date have noted an increase in prevalence of autism among males over females (e.g. Gillberg, 1984; Lotter, 1966; Treffert, 1970). Most estimates of the ratio of males to females are between 1.4:1 and 3.4:1. Some researchers have also indicated that the male to female ratio increase with IQ, with significantly more males with autism than females having an IQ greater than 50 (Lotter, 1966; Wing, 1981).

Etiology of Autism

The etiology of autism remains one of the primary controversies in the field of mental health, with psychogenic, genetic, and neurochemical factors being the most common explanations. The initial hypotheses regarding the origins of autism focused on parental pathology. In Kanner's (1943) earliest publication on infantile autism he made several observations regarding the parents of the autistic children. Kanner noted that all of the children came from highly educated parents who were preoccupied by their careers and cold and formal

in their personal lives. Of all 11 families, Kanner stated that there were very few warm-hearted parents. While Kanner was careful to state that autism could not be entirely attributed to parent relations, he planted the seed for further psychogenic explanations of autism, which included theories on inadequate emotional connections between mother and child (Mahler, 1952; Bettelheim, 1967) and possessiveness on the part of the mother (Bettelheim, 1967). Without research to support them, these theories eventually receded and were replaced by genetic and physiological hypotheses (Rimland, 1964).

Genetic factors contributing to autism are well documented. Folstein and Rutter (1977) examined 11 pairs of monozygotic twins and 10 pairs of dizygotic twins, with one child in each pairing having a diagnosis of autism. Folstein and Rutter found a 36% pair wise concordance rate for autism in the monozygotic twins, as opposed to 0% for the dizygotic twins. Results of another twin study obtained similar results. Ritvo et al. (1985) studied 23 pairs of monozygotic and 17 pairs of dizygotic twins and found a 95.7% concordance rate among the monozygotic twins as opposed to a 23.5% rate among the dizygotic twins. These results must be viewed cautiously due to possible selection bias and lack of random sampling in these studies (Pauls, 1987; Phelps & Grabowski, 1991).

Finally, research has focused on neurochemical abnormalities as an etiological factor in autism. Several neurotransmitters have been posited to play a role in the expression of autism. Serotonin has been the most widely studied neurotransmitter in autism research (Volkmar & Anderson, 1989). A study by Anderson and Hoshino (1987) found the group mean blood serotonin level in autistic subjects to be 17% to 128% higher than controls. While these results have been replicated, it is important to note that the serotonin levels in the autistic sample

appeared to be normally distributed (Yuwiler, Geller, & Ritvo, 1985). In addition, the mechanism for this possible serotonin elevation remains unclear.

Adult Outcome of Autism

As first described, autism was called infantile autism, and only studied in children (Kanner, 1943). In the 60 years since Kanner, researchers have remained focused on the childhood expression of the disorder, while at the same time children with autism are growing into autistic adults. Given the vast literature on childhood autism, it is remarkable how little consideration has been paid to the characteristics, abilities, and needs of these adults.

The little research known about the prognosis for adults with autism is bleak (Kanner, 1971). Investigators in several long-term follow up studies found that at least two thirds of individuals with autism were unable to live independently, and approximately half of individuals with autism lived in residential facilities (Kanner, 1971; Lotter, 1974). However, this figure is an improvement from only a few years earlier when nearly 95% of people with autism lived in state run institutions (Hitzing, 1987).

A possible reason for the institutionalization of adults with autism is the presence of maladaptive behaviors in the population. Researchers have shown that aggression, self-injury, and stereotypy are most problematic and common in autism (Gardner & Cole, 1990; Fee & Matson, 1992; Rojahn & Sisson, 1990). It is thought that these maladaptive behaviors interfere with learning both social and daily living skills. However, the variables that maintain these problem behaviors in adults with autism are unknown (Dawson, Matson, & Cherry, 1998).

The most pervasive problem facing individuals with autism into adulthood is social impairment. A study on peer relations found that 89% of adults with autism lacked an identifiable friend and 56% presented a flat affect when relating to other people (Rumsey,

Rapaport, & Sceery, 1985). Even the highest functioning adults with autism continue to evince difficulties with social relationships and communication (APA, 1994). Having discussed the adult outcome of autism, the next topic requiring consideration is the behavioral assessment of autism.

Behavioral Assessment of Autism

Previous sections of this paper discussed the history, characteristics, prevalence, and etiology of mental retardation and autism, as well as the adult outcome of the disorder. The next area of examination is the behavioral assessment of autism. According to Powers (1988), the assessment of individuals with developmental disabilities centers on three fundamental goals. The first goal is to produce observation and evaluation of behavior using both traditional instruments, such as intelligence tests, as well as instruments designed expressly for use on the specific population. The second goal of assessment is to conduct data analysis in order to provide functional evaluation of strengths and weaknesses across various behavioral domains. The third goal of assessment is to synthesize behavioral, assessment and diagnostic data to guide production of intervention and treatment plans.

This study concentrated on Powers' first goal of assessment, the use of instruments designed for use on an autistic population. The following instruments will be described and evaluated: the Diagnostic Checklist for Behavior-Disturbed Children (Diagnostic Checklist), Autism Behavior Checklist (ABC), Behavior Rating Instrument for Autistic and Atypical Children (BRIAAC), Behavioral Observation System (BOS), and the DASH-II. The instrument used in the present study, the Childhood Autism Rating Scale (CARS) will be discussed last.

Rimland (1964) developed one of the first behavior rating scales. Based upon the core symptoms defined by Kanner (1943), the Diagnostic Checklist contained items regarding social interaction, speech, reaction to stimuli, intelligence, family information, and psychological development. After initial study, parent report suggested that autistic behaviors became more idiosyncratic and core symptoms more diffuse after the age of five (Rimland, 1968). As a result, the Diagnostic Checklist later contained items that applied only to children under the age of five

years (Rimland, 1971). Several methodological problems limit the use of the Diagnostic Checklist. Namely, Rimland failed to provide objective definitions of terms (Masters & Miller, 1970) and relied solely on parent recall of behavior. In addition, the Diagnostic Checklist has not been tested for interrater reliability and is unable to distinguish between autistic and schizophrenic children (DeMyer, Churchill, Pontius, & Gilkey, 1971).

The ABC consists of behaviors taken from several sources, which include the Diagnostic Checklist (Rimland, 1964), Creak's (1964) criteria, and Kanner's (1943) criteria. The ABC contains 57 items organized into five symptom groupings: sensory, relating, body and object use, language, and social and self-help skills. Raters indicate the presence or absence of each behavior, and each item is then weighted and compared with standardized groups arranged by chronological age. Advantages of the ABC are the use of teachers as opposed to parents as raters, the simple and straightforward scoring, and the acknowledgement of the effect of age on the disorder with the inclusion of separate profile charts for different age groups (Volkmar, Cicchetti, et al., 1988). One test of the ABC's validity found that it correctly predicted autistic group membership in 100% of individuals and trainable mentally retarded group membership in 95% of individuals (Teal & Wiebe, 1986). Initial reports of interrater reliability and split-half reliability were high, but based on observation of only fourteen children (Krug et al., 1980). Criticisms of the ABC are the focus on aberrant symptomology, lack of measurement of prosocial behavior, and the need for a revised cutoff score (Sevin, Matson, Coe, Fee, & Sevin, 1991).

Another instrument developed for the assessment of autism is the BRIAAC (Ruttenberg, Dratman, Frankno, & Wenar, 1966; Ruttenberg, Kalish, Wenar, & Wolf, 1977). This test consists of eight scales: relationship to adults, communication, drive for mastery, vocalization

and expressive speech, sound and speech reception, social responsiveness, body movement, and psychobiological development. The behavioral descriptions in the BRIAAC were all empirically derived from clinical notes. Scores are based on clinical observation. While the authors demonstrated that the BRIAAC could distinguish between normal, mentally retarded, and autistic children (Wolf, Wenar, & Ruttenberg, 1972), it cannot differentiate between autistic children and those with other disorders (Cohen et al., 1978).

Another important measure is the BOS, which was designed to differentiate between autistic and nonautistic children and to provide a sound basis for the description of autism in research (Freeman, Ritvo, & Schroth, 1984). The BOS was comprised of 24 items divided into four groups: solitary, relationship to objects, relationship to people, and language. Interrater reliability of the BOS is high, with a mean item correlation coefficient of 0.71 (Freeman et al., 1984). However, the BOS does not provide a diagnostic cutoff score and fails to discriminate between an autism diagnosis and other childhood psychopathologies.

As previously described, the DASH-II is an instrument designed for the broad screening of psychopathology in individuals with severe and profound mental retardation (Matson, 1995b; Matson, Gardner, Coe & Sovner, 1991). The PDD/ autism subscale of the DASH-II has demonstrated acceptable internal consistency, strong predictive ability, and is significantly correlated to the CARS (Matson, Smiroldo & Hastings, 1998).

Childhood Autism Rating Scale

The instrument under investigation in the current study is the CARS. The CARS was chosen for this study because it is considered the gold standard in the field (Matson et al., 1998; Morgan, 1988; Sturmey & Sevin, 1994). The CARS was designed to differentiate between autistic and other developmentally disordered children. Development of the CARS began in 1966 with the production of a scale that incorporated the criteria of Kanner (1943) and Creak (1964), and characteristic symptoms of childhood autism (Schopler et al., 1980).

The 15 items of the CARS are (Schopler, Reichler, & Renner, 1988): 1) Relating to people; 2) Imitation; 3) Emotional response; 4) Body use; 5) Object use; 6) Adaptation to change; 7) Visual response; 8) Listening response; 9) Taste, smell, and touch response and use; 10) Fear or nervousness; 11) Verbal communication; 12) Nonverbal communication; 13) Activity level; 14) Level and consistency of intellectual response; 15) General impressions.

Each scale is rated with a score of 1 (with normal for child's age), 2 (mildly abnormal), 3 (moderately abnormal), or 4 (severely abnormal). Midpoint scores of 1.5, 2.5, and 3.5 are also possible. Total CARS scores range from 15 to 60, with a score of 30 serving as the cutoff for a diagnosis of autism.

Initial psychometrics for the CARS were determined using 537 children enrolled in the Treatment and Education of Autistic and related Communication handicapped Children (TEACCH) program over a 10-year period (Schopler et al., 1980). Fifty-one percent of the children studied scored above the cutoff score of 30. Schopler et al. (1980) observed the existence of a bimodal distribution among these scores, leading them to develop criteria to differentiate between those with mild to moderate autism and those with severe autism. Children

those score exceeded 36 and who received a rating of three or greater on at least five subscales were categorized as being severely autistic.

Internal consistency of the CARS was high, with a coefficient alpha of .94 (Schopler et al., 1988), indicating the degree to which all of the 15-scale scores constitute a unitary phenomenon rather than several individual behaviors. Interrater reliability was established using two raters for 280 cases. The average reliability of .71 indicated good overall agreement between raters. Twelve-month test-retest data was also collected and found that the means were not significantly different from the first testing to the second. Criterion-related validity was determined by comparing CARS diagnoses to diagnoses made independently by child psychologists and psychiatrists. Diagnoses correlated at $r = .80$, which indicated that the CARS diagnosis was in agreement with clinical judgments. The CARS has also been shown to have 100% predictive accuracy when distinguishing between groups of autistic and mentally retarded children, which was superior to the ABC and Diagnostic Checklist (Teal & Wiebe, 1986).

Validity of the CARS under different settings is of particular importance to the present study. CARS scores of 41 children taken through parent interview were compared to scores derived from direct observation. Mean scores under the two conditions were not significantly different and the correlation of $r = .83$ further indicated good agreement. In addition, diagnoses based on parent interview and direct observation agreed in 90% of the cases. The authors suggest that valid CARS ratings and diagnoses can be achieved through parent interview (Schopler et al., 1988).

Of the autism rating scales discussed, the CARS is the only scale used in research with adolescents or adults. Mesibov, Schopler, Schaffer, and Michal (1989) examined the diagnostic ability of the CARS in adolescents and adults with autism. This study compared the CARS

scores of 89 individuals before the age of 10, (mean age of 8.7 years,) with their scores after the age of 13, (mean age of 15.9 years.) Fifty-nine of the original 89 participants were diagnosed with autism before the age of ten, with a mean CARS score of 38.47. For those 59 participants, adolescent CARS scores revealed a significant decrease, with a mean score of 35.54. Significant improvement in adolescent scores, (i.e. significant decreases in abnormal behavior,) were found on the imitation, body use, object use, adaptation to change, listening response, sensory response and use, verbal communication, nonverbal communication, and activity level scales. The only score with a significant increase over time was the general impression scale. It was suggested that a cutoff score of 27, as opposed to 30, be used when administering the CARS to adolescents and adults. This recommendation stemmed from the observation that when the cutoff score was lowered by three points, (corresponding to the mean difference in scores between the two age groups,) the percentage of individuals accurately diagnosed as autistic before age 10 and after age 13 jumped from 81% to 92%. Overall, Mesibov et al. (1989) suggested that the CARS is a good screening instrument for adolescents and adults.

While Mesibov et al.'s (1989) study provided valuable insight into the course of the disorder over time; the generalization of the use of the CARS on adults based on its results is premature. The mean age of participants in the "adolescent and adult" age group was only 15.9 years. As a result, the study failed to demonstrate the ability of the CARS to diagnosis autism in adults, (i.e. individuals over 21 years old.) The CARS has been used with adults in clinical settings, but researchers have yet to verify its diagnostic ability with this population.

Interrater Agreement and Staff Accuracy

An issue of concern in assessment is the degree of interrater agreement in rating scales. In the assessment of individuals with behavior problems, it is recommended that information be compiled from multiple sources (Haynes & O'Brien, 2000). Literature is available on the agreement between various groups of people, including the agreement between teachers, parents, staff, and clinicians.

Achenbach, McConaughy and Howell (1987) performed a meta-analysis of 119 studies that compared informant's ratings of behavioral and emotional problems in children. The highest correlations were found in situations where informant pairs had similar relationships with the subject, (such as mothers and fathers, teachers and teacher aides, clinicians and dormitory personnel in residential settings.) The mean Pearson *rs* ranged from .54 for pairs of mental health workers to .64 for pairs of teachers. The overall mean agreement was .60. When informants with different relationships to the subject, (such as parents and teachers, parents and mental health workers, etc.,) were paired the correlations were considerably lower. The mean correlations ranged from only .24 for parents paired with mental health workers, to .42 for teachers paired with trained observers.

Achenbach et al.'s (1987) results demonstrate the poor interrater agreement typically found in the assessment of childhood behavior problems. The lower rates of agreement found in pairs of informants with different relationships to the subject indicate that the type of informant has an effect on agreement. Achenbach et al. (1987) do not believe that the low correlations necessarily represent poor reliability. The authors suggest that different types of informants function in different settings and may interact

differently with the subject. These results further emphasize the need for multiaxial assessment.

In the ideal situation, an individual suspected of having a behavioral problem or developmental disability would be observed by a psychologist with extensive training and education in the problem area. These observations would take place in multiple settings and over several periods of time. This approach is not practical in most institutions, due to the large caseloads carried by psychologists and the cost of the psychologist's time. As a result, the institution's direct-care staff is relied upon for a great deal of information on the clients.

Direct-care staff in most residential facilities varies to a great extent in terms of education, experience, length of employment, attitude, and motivation. Thus, the concordance between staff members is highly variable, in addition to the unknown agreement between staff members and psychologists. Several researchers have investigated the variability of staff report in residential settings. McGill et al. (2001) examined staff reports of three challenging behaviors: stereotypy, self-injury, and aggression. Four staff members were interviewed for each of the 22 participants. McGill et al. (2001) found the average correlation between pairs of raters to be .66 for total scores, .69 for self-injury, .73 for aggression, and .52 for stereotyped behaviors. The study's authors questioned whether the high degree of variability between raters reflected a lack of reliability or a real variability in the challenging behaviors of the individuals with developmental disabilities.

The accuracy of direct-care staff reports has also come under criticism in a number of studies. Green et al. (1988) compared the reinforcer preferences of seven

individuals with profound mental retardation with staff opinion. Reinforcer preference was determined by using a systematic assessment of approach and avoidance behaviors. Staff opinion was assessed with a survey of reinforcers rated on a Likert scale of 1 (least preferred) to 5 (most preferred). Results indicated that there was no significant correlation between individuals' preferences based on the systematic assessment and staff opinion.

Parsons and Reid (1990) also studied direct-care staff accuracy. In this study, clients' food and drink preferences were assessed using a structured behavioral assessment and a traditional staff survey. Similar to Green et al. (1988), Parsons and Reid (1990) found that staff opinion of clients' preferences did not concur with results of the structured assessment. These studies reflect that staff may not always provide accurate accounts of client preferences or other behaviors.

The goal of this study was to address the issues of interrater agreement and accuracy of direct-care staff report by measuring the degree of reliability among staff members, as well as the concordance between staff and trained, masters-level mental health professionals. This information is vital to our understanding of behavioral assessment instruments, such as the CARS, and our general understanding of the way assessments are conducted in institutions.

Autism and Mental Retardation

The present study focused on the behavioral assessment of adults dually diagnosed with autism and mental retardation. In the first description of autism, Kanner (1943) reported that children with autism possessed near normal or above average intelligence. However, researchers have shown a strong relationship between autism and mental retardation (Janicki & Jacobson, 1983; Schopler et al., 1980; Wing & Gould, 1979). A descriptive study of 314 autistic adults in the state of New York (Janicki & Jacobson, 1983) found that 85% also had a diagnosis of mental retardation. Of these, 43% of the people were profound, 30% severe, 17% moderate, and 10% mild. Wing and Gould (1979) examined an entire geographical area for children under the age of 15 with severe mental retardation. Wing and Gould found that 11% of the children were autistic, 40% were socially impaired but not autistic, and 49% had age-appropriate social skills.

In addition to investigating the prevalence of the two disorders, researchers have also looked at the outcome of the dually diagnosed population. A study conducted by Schopler et al. (1980) on the co-existence of mental retardation and autism in children found that the degree of mental retardation was positively correlated to the degree of autism. Thus, the children with profound mental retardation were the most severely autistic. None of the severely autistic children in the sample functioned within the normal range of intelligence. Schopler et al. (1980) suggested the possibility that the classically autistic child with normal intelligence described by Kanner was not the severely autistic child seen today. Instead, Kanner's classical autism corresponds with mild-to-moderate autism as seen in Schopler et al.'s (1980) subject pool.

Intelligence appears to be the strongest predictor of outcome in autism. An IQ above 60 is almost always associated with a fair to good global outcome rating, whereas people with IQs below 60 are most representative of those with poor outcomes (Rutter, Greenfield, & Lockyer,

1967). Children with autism and comorbid mental retardation may have different deficits than those with autism and normal intelligence. While autistic individuals with high and low IQs both exhibit the core features of autism, (i.e. social and language deficits, insistence on sameness, and abnormal response to sensory stimuli,) those with mental retardation show more severe social deficits and are more likely to evince deviant social responses (Rutter, 1978). Patterns of cognitive deficits also differ in autistic individuals with mental retardation and those with normal intelligence. Generally, individuals with autism and mental retardation show greater cognitive deficits, specifically in the areas of sequencing and feature extraction (Hermelin & O'Conner, 1970). Bartak and Rutter (1976) compared the mathematical abilities of children with autism with IQs above 70 to those with IQs below 70. Three-fourths of children with autism and an IQ above 70 acquired competence in basic arithmetic, compared to less than 1/5 of autistic children with an IQ below 70.

In addition, the DSM-IV addressed the issue of the comorbidity of mental retardation and autism (APA, 1994). The DSM-IV definition of mental retardation is more common in females with autism over males. In addition, the DSM-IV recommended that when determining whether an individual with mental retardation also necessitates a diagnosis of autism, the dual diagnosis should be reserved for situations in which the specific behavioral characteristics of autism (i.e. deficits in social and communicative skills,) are present.

Several researchers have attempted to address the lack of information on the overlap of the behavioral characteristics of autism and mental retardation, but not without limitations. In 1986 Lund examined the behavioral symptoms of 302 adults with mental retardation in Denmark, reporting that 27% of the sample displayed social withdrawal, 28% had abnormal language, and 22% had stereotyped behavior. However, only 7.6% met full criteria for autism.

Several problems existed within this study, namely the inclusion of only broad categories of symptoms, the lack of a specific diagnostic system for assessing the presence of autism, and the small number of individuals with severe and profound mental retardation ($n = 59$).

A study by Shah, Holmes, and Wing (1982) reported on the presence of severe social impairment and features of autism in a large sample of residents in a long-stay mental handicap hospital in London ($n = 761$). Participants were assessed using the Disability Assessment Schedule (Holmes, Shah, & Wing, 1982), which included items regarding social interaction, speech abnormalities, imaginative activities, repetitive routines, stereotypies, and behavior problems. Shah et al. found that 4% of their sample was autistic, 34% was socially impaired but not autistic, and 62% was “sociable”. Unfortunately, this study failed to report the level of mental retardation of its participants, which greatly limits its generalization.

Finally, Janicki and Jacobson (1983) investigated the clinical features of a sample of 314 adults with autism. Data regarding participants’ ages, IQ, and adaptive abilities were compiled from a statewide database. The participants in this study were young, only 24.2% were over the age of 21. Eighty-five percent of the participants were classified as mentally retarded, with 43% functioning within the profound range, 30% severe, 17% moderate, and 10% mild. This study was limited by the lack of information regarding the method of diagnosing autism, the specific behavioral characteristics of the autistic population as a whole, and those of the mentally retarded subgroup, and the young age of the participants. The three studies discussed here, along with the scarcity of information regarding the ability of the current diagnostic systems to reliably assess autism in adults with mental retardation, provided the foundation for the current study.

Some researchers have questioned the need for differential diagnosis in individuals with mental retardation (Powers & Handleman, 1984). Fortunately, the field has progressed passed

the notion that differential labeling is arbitrary, and now recognize the benefits of diagnostic accuracy (Parks, 1988). The first benefit of accurate diagnosis of autism in individuals with autism is more effective treatment planning. Persons with autism exhibit quantitative and qualitative differences in social skills and language that differentiate them from individuals with mental retardation alone (Njardvik et al., 1999). As a result, better diagnosis of autism in this population will lead to use of treatments that have support for use with individuals with autism, such as increased environmental structuring and specific supports for transitioning between residential and community settings (Parks, 1988). Using the appropriate individualized treatment may result in providing an environment with the least restrictions, which is a fundamental goal of residential facilities (Thomas, 1994). It is also possible that clients' global functioning is being underestimated due to masking by autistic symptoms. The second benefit of accurate diagnosis is a clarity of mission for the staff. Individuals with autism present specific challenges that may be difficult or frustrating for those who do not understand the condition. By improving our ability to diagnose autism in individuals with mental retardation, more effective treatments can be prepared and direct-care staff can be trained to deal with individuals' strengths and weaknesses more effectively (Thomas, 1994).

In closing, this introduction provided background into the history, prevalence, etiology, and diagnosis of mental retardation and autism, as well as the field of dual diagnosis in mental retardation. The dearth of research on the assessment of autism in adults is readily apparent. Despite the expressed need for research on this population (Cohen, Paul, & Volkmar, 1986; Volkmar & Schwab-Stone, 1996), clinicians still have no empirically supported assessment instruments to aid in their diagnostic decision making when faced with an adolescent or adult

suspected of having autism. Research such as the current project attempts to address this concern in order to improve the quality of life of these individuals.

Purpose

Autism is a disorder that affects individuals across the lifespan. However, most of research involves children. As a result, minimal literature is available regarding the use of diagnostic screening tools on adults with autism. With the notion that as many as 85% of individuals with autism have mental retardation (Janicki & Jacobson, 1983), adults with comorbid autism and mental retardation have been neglected.

In order to address this lack of research and thus appropriate clinical practice relative to adults with autism, the present study was designed to provide the first reliability and concordance data in adults with mental retardation on two commonly used methods for diagnosing autism, the Childhood Autism Rating Scale (CARS) and the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV). The CARS was selected based on its well-researched psychometrics in children (Mesibov et al., 1989; Schopler et al., 1980; Schopler et al., 1988; Sevin et al., 1991) and its status as the gold standard of autism rating scales (Matson et al., 1998; Morgan, 1988; Sturmey & Sevin, 1994). The DSM-IV is commonly used in clinical settings for diagnosing autism and other developmental and emotional disorders and is tied to insurance funding. Thus, studying these assessment methods for adults with autism has both theoretical and practical significance.

More specifically, the current study's purpose had three goals. The first goal was to begin establishing reliable methods for identifying autism in adults by examining the test-retest and interrater reliability of the CARS and DSM-IV on adults with severe and profound mental retardation. The current study addressed this issue, while also evaluating the appropriateness of the use of the CARS and DSM-IV with direct-care staff serving as informants. A comparison between direct-care staff and masters-level mental health professionals was made to determine

whether individuals with varying educational backgrounds and experience in the field make similar judgments regarding presence of autism using these two instruments. Researchers have shown the CARS to be valid when completed by parents as compared to direct observation by trained clinicians (Schopler et al., 1988). Therefore, the current study assessed whether this same concordance occurred when direct-care staff and trained professionals provided ratings.

The second goal of this study was to investigate the concordance between CARS and DSM-IV “diagnoses” as determined by both direct-care staff and masters-level mental health professionals. These two instruments are purported to measure the same construct, autism. If these instruments do measure the same construct, then one would expect their diagnoses to agree. Understanding of the concordance between instruments is a necessary component to effective assessment when clinicians utilize multiple diagnostic tools.

The final goal of the study was to provide a preliminary examination of the appropriateness of the use of the CARS and DSM-IV on individuals with profound mental retardation (PMR; n = 46) and severe mental retardation (SMR; n = 46). The purpose of comparing these two groups was to determine whether these two instruments are sensitive to the qualitative differences between autism and mental retardation in adults. Many of the characteristics of autism, including social and language deficits, repetitive language and stereotypic movements, are also common in individuals with mental retardation (Janicki & Jacobson, 1983; Lund, 1986; Shah et al., 1982). As severity of mental retardation increases, so does the frequency and intensity of these behaviors. Part of determining the utility of the CARS and DSM-IV in this population is establishing whether these instruments can differentiate between characteristics of mental retardation and autism.

The three goals under examination in this study provide the first reliability and concordance data on the CARS and DSM-IV in adults. Autism has been understudied in adults, leaving clinicians with no empirically supported instruments with which to diagnose autism in this population. It is our contention that the need for such tools exists for this population, due to the only recent attention paid to dual diagnosis in mental retardation. Many adults currently in residential and community facilities were not properly diagnosed as children, a state of affairs that will continue until instruments are developed or renormed with this population. Thus, research such as this will provide a basis for more accurate diagnoses, better tailored treatment planning, and more comprehensive staff training. Additionally, we currently do not have the evidence to support the lack of need for more specific diagnoses in this population. Research into the assessment of autism in this population will provide clinicians with the instruments necessary to provide adequate assessment and treatment.

Method

Participants

Participants were residents at Pinecrest Developmental Center (PDC). PDC is a state-run facility in Louisiana that is home to 650 individuals. The individuals residing at PDC represent various ages, gender, race, and levels of mental retardation. Approval from the Institutional Review Board (IRB) was obtained.

Fifty individuals with profound mental retardation (PMR) and 50 with severe mental retardation (SMR) were originally randomly selected for inclusion in this study. During the course of data collection, two individuals from the PMR group and four from the SMR group were removed from data collection due to death or transfer to a community group home. In order to maintain equal numbers between groups, one additional member of the PMR group was randomly dropped from the study, leaving 46 individuals in the PMR group and 46 in the SMR group.

The PMR group consisted of 33 females and 13 males. Thirty-three subjects in the PMR group were Caucasian and 13 were African American. The average age of the PMR group was 47.87 years. The SMR group consisted of 21 females and 25 males. Forty-four of the SMR participants were Caucasian, one was African-American, and one was Native American. The average age of the SMR group was 52.11 years. A two-tailed t-test used to compare the average ages of the PMR ($M = 47.87$) and SMR ($M = 52.11$) did not find a significant difference ($t(90) = -1.30$, ns).

Dependent Variable

Childhood Autism Rating Scale (CARS). The CARS is a 15-item informant and observation based rating scale (Schopler et al., 1988). The items on the CARS are as follows: 1)

relating to people; 2) imitation; 3) emotional response; 4) body use; 5) object use; 6) adaptation to change; 7) visual response; 8) listening response; 9) taste, smell, and touch response; 10) fear or nervousness; 11) verbal communication; 12) nonverbal communication; 13) activity level; 14) level and consistency of intellectual response; 15) general impression. Each item is scored on a 7-point likert scale in which 1 = within normal limits for that age, 1.5 = very mildly abnormal for that age, 2 = mildly abnormal for that age, 2.5 = mildly-to-moderately abnormal for that age, 3 = moderately abnormal for that age, 3.5 = moderately-to-severely abnormal for that age, and 4 = severely abnormal for that age. The CARS has high internal consistency, good interrater reliability, and high criterion-related validity for children and adolescents (Schopler et al., 1988).

DSM-IV Checklist. The DSM-IV diagnosis of autistic disorder is divided into three categories. Category one includes the following four items on social interaction; 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; 4) lack of social or emotional reciprocity. Category two includes the following four items regarding language; 1) 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); 2) in individuals with adequate speech, marked impairment in the ability to initiate or sustain a conversation with others; 3) stereotyped and repetitive use of language or idiosyncratic language; 4) lack of varied, spontaneous make-believe play or social imitative play appropriate to developmental level. The third category is comprised of the following four items pertaining to repetitive and stereotyped patterns of interest; 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; 4) persistent preoccupation with parts of objects. Each item was rated as either present (a score of one) or not present (a score of zero) in each individual. In order to meet a diagnosis of autism, an individual must present with a total of six or more endorsements, with at least two coming from category one, and at least one coming from both categories two and three. The DSM-IV also states that the presence of abnormal functioning in either/or social interaction, language, and symbolic or imaginative play must be present in the individual prior to the age of three. Due to the difficulty in acquiring early childhood history of the clients in a residential facility, and the current study's purpose of evaluating endorsements of DSM-IV criteria rather than making diagnoses, this criterion was not included in the DSM-IV Checklist. The DSM-IV Checklist can be found in Appendix A.

Procedure

The CARS and DSM-IV checklist were administered for each of the 92 participants using two primary direct-care staff members who had known the individual for at least six months prior to the study. Upon hire, all direct-care staff attended a 40-hour training seminar that provided training on behavior problems, treatments, and care of individuals with mental retardation and developmental disabilities. In addition to assisting with daily living, direct-care staff is responsible for data collection on problem behaviors, treatment plan implementation, and completing informant-based rating scales for residents' early psychological evaluations. Test-retest data was collected two to four weeks after the initial interview with one direct-care staff for each participant.

Bachelors and masters-level college graduates served as interviewers. The interviewers familiarized themselves with the CARS administration manual (Schopler et al., 1988) and individual CARS and DSM-IV items. Interviews took place in a quiet area of the participants' homes. Interviewers followed the instructions in Appendix B. Each item was read to the direct-care staff verbatim. Upon completion of the item, the interviewer instructed the direct-care staff to provide a rating.

In addition, one masters-level mental health professional completed a CARS and a DSM-IV checklist for 84 of the participants. These individuals each held a masters degree in psychology. Master-level mental health professionals are responsible for conducting early psychological evaluations, identifying psychopathology, assessing behavior problems, designing treatment plans, and training direct-care staff on the implementation of these plans. Interviews with the master-level mental health professionals were conducted in their offices using the same protocol as that used for direct-care staff interviews. Master-level mental health professionals were not available for 10 of the participants. All informants were blind to the purpose of the study.

Results

Reliability

Test-Retest Reliability. The total CARS score for each participant was calculated and compared for repeated direct-care staff ratings using a Pearson product moment correlation (r) due to the continuous nature of the variable (Hinkle, Wiersma, & Jurs, 1998). Test-retest reliability of the CARS with direct-care staff was $r = .86$, $p < .01$. Test-retest reliability of the individual CARS items is provided in Table 1. The correlation coefficients of the individual items ranged from .42 for Level and Consistency of Intellectual Response to .85 for Verbal Communication. All correlations coefficients could be classified as having fair to excellent clinical significance (Cicchetti, 1994).

DSM-IV diagnoses were made for each participant using DSM-IV scoring criteria (APA, 1994). Test-retest reliability of DSM-IV diagnoses was assessed using the kappa correlation coefficient (κ) in order to measure agreement between paired, dichotomous variables while removing chance agreement (Hinkle, Wiersma, & Jurs, 1998). Kappa is the most conservative means of calculating reliability coefficients in dichotomous variables (Hinkle, Wiersma, & Jurs, 1998). Test-retest reliability was found to be $\kappa = .64$, $p < .001$. Test-retest reliability of the 12 DSM-IV Checklist items was also determined in order to provide a comprehensive representation of reliability. Table 2 presents the test-retest reliability of the DSM-IV Checklist items. Individual item kappa correlations ranged from poor for Lack of Make-Believe play ($\kappa = .15$) to excellent for Stereotyped Motor Mannerisms ($\kappa = .82$).

Interrater Reliability. Interrater reliability was examined for both the CARS and DSM-IV using two pairs of informants: agreement between two direct-care staff members and agreement between direct-care staff and the masters-level mental health professionals. Total

Table 1

Test-Retest Reliability of CARS Items

Item	<i>r</i>
Relating to People	.64*
Imitation	.59*
Emotional Response	.43*
Body Use	.66*
Object Use	.70*
Adaptation to Change	.56*
Visual Response	.60*
Listening Response	.51*
Taste, Smell, and Touch Response and Use	.64*
Fear or Nervousness	.54*
Verbal Communication	.85*
Nonverbal Communication	.46*
Activity Level	.45*
Level and Consistency of Intellectual Response	.42*
General Impressions	.60*

* $p < .01$

Table 2

Test-Retest Reliability of DSM-IV Checklist Items

Item ^a	κ
Marked Impairment in Nonverbal Behaviors	.31*
Failure to Develop Peer Relations	.43**
Lack of Spontaneous Sharing	.40**
Lack of Social or Emotional Reciprocity	.45**
Delay in the Development of Spoken Language	.62**
Marked Impairment in Initiation of Conversation	.47**
Stereotyped Language	.45**
Lack of Make-Believe Play	.15
Stereotyped and Restricted Interests	.32*
Inflexible Adherence to Nonfunctional Routines or Rituals	.26*
Stereotyped Motor Mannerism	.82**
Preoccupation with Parts of Objects	.31*

^a see Appendix B for complete description of items

** p < .001

* p < .01

CARS scores were compared between raters using a Pearson r correlation coefficient (Hinkle, Wiersma, & Jurs, 1998). The correlation between the two direct-care staff members on the CARS was $r = .42$, $p < .01$. The correlation between direct-care staff and the masters-level mental health professionals on total CARS scores was $r = .50$, $p < .01$. Table 3 presents the interrater reliability of the individual CARS items for both pairs of informants. For pairs of direct-care staff, interrater reliability of the items was below $r = .50$ for all but two items: Verbal Communication and Nonverbal Communication. For direct-care staff/professional pairs, all correlations fell below $r = .40$ except for Verbal Communication, which had an r of .81.

For the DSM-IV diagnoses, a kappa correlation coefficient was used to assess interrater reliability because the variables were dichotomous and kappa is the most conservative coefficient available in this situation (Hinkle, Wiersma, & Jurs, 1998). Interrater reliability for direct-care staff pairs was $\kappa = .32$, $p < .01$ and $\kappa = .31$, $p < .01$ for pairs of direct-care and professionals. The individual item interrater reliability between direct-care staff members, as well as the interrater reliability between direct-care staff and masters-level mental health professionals are shown in Table 4. For direct-care staff pairs, the only good correlation found was for Delay in the Development of Spoken Language; the remaining correlations were all below $r = .40$ and classified as poor (Cicchetti, 1994). A similar result was found when direct-care staff and professionals were paired. In this case, Delay in the Development of Spoken Language ($r = .47$) was the highest correlation found.

Concordance

Agreement between CARS and DSM-IV diagnoses was evaluated for the direct-care staff, (using the first direct-care staff rating,) as well as for the masters-level mental health professionals. CARS scores were converted to diagnoses using the cutoff score of 30 (Schopler

Table 3

Interrater Reliability of CARS Items

Item	Direct-Care	Direct –Care and Professionals
	<i>r</i>	<i>r</i>
Relating to People	.47 **	.33 **
Imitation	.36 **	.18
Emotional Response	.21 *	.19
Body Use	.20	.29 **
Object Use	.45 **	.32 **
Adaptation to Change	.29 **	.14
Visual Response	.15	.24 *
Listening Response	.23 *	.33 **
Taste, Smell, and Touch Response and Use	-.03	.19
Fear or Nervousness	.34 **	.10
Verbal Communication	.65 **	.81 **
Nonverbal Communication	.52 **	.33 **
Activity Level	.50 **	.21
Level and Consistency of Intellectual Response	.25 *	.08
General Impressions	.27 **	.33 **

** p < .01

* p < .05

Table 4

Interrater Reliability of DSM-IV Checklist Items

Item ^a	Direct-Care	Direct –Care and Professionals
	κ	κ
Marked Impairment in Nonverbal Behaviors	.35***	.15
Failure to Develop Peer Relations	.30**	.19
Lack of Spontaneous Sharing	.20	.40***
Lack of Social or Emotional Reciprocity	.36***	.28**
Delay in the Development of Spoken Language	.65***	.47***
Marked Impairment in Initiation of Conversation	.35***	.13
Stereotyped Language	.27**	.14
Lack of Make-Believe Play	.14	.16
Stereotyped and Restricted Interests	.14	.02
Inflexible Adherence to Nonfunctional Routines or Rituals	.13	.03
Stereotyped Motor Mannerism	.23*	.28**
Preoccupation with Parts of Objects	-.08	.14

^a see Appendix B for complete description of items

*** p < .001

** p < .01

* p < .05

et al., 1988). All participants who scored below 30 did not meet the criteria for autism, and all who scored 30 or above were considered autistic. Due to the independent and dichotomous nature of the variables, a chi-square test was used to assess agreement while controlling for agreement due to chance. For direct-care staff, the agreement between CARS and DSM-IV diagnoses was significant ($\chi^2(1) = 20.70$, $p < .001$). As shown in the Table 5, the CARS and DSM-IV agreed on diagnosis 71 of the 92 participants. For the masters-level mental health professionals, the concordance between CARS and DSM-IV diagnoses was also significant ($\chi^2(1) = 36.47$, $p < .001$). Table 6 shows that the CARS and DSM-IV agreed on diagnosis for all but 9 of the 84 participants when comparing the ratings of masters-level mental health professionals.

Table 5

Chi-Square for Direct-Care Staff CARS and DSM-IV Checklist Diagnostic Agreement

CARS		
DSM-IV	Autism	No Autism
Autism	17	6
No Autism	15	54

Group Comparison

Individuals with profound mental retardation (PMR) and severe mental retardation (SMR) were compared to determine whether a difference existed between groups in total CARS score and number of autism diagnoses on the DSM-IV. An independent, two-tailed t-test was

Table 6

Chi-Square for Professionals CARS and DSM-IV Checklist Diagnostic Agreement

CARS		
DSM-IV	Autism	No Autism
Autism	11	2
No Autism	7	64

conducted to compare the CARS scores of the PMR and SMR groups (Hinkle, Wiersma, & Jurs, 1998). A two-tailed t-test was chosen due to the exploratory nature of this research and the lack of knowledge about the direction of the relationship. A significant difference was found between the PMR ($M = 31.97$) and SMR groups ($M = 22.44$), $t(90) = 6.68$, $p < .001$. The PMR group had a significantly higher average CARS score than the SMR group.

For the DSM-IV diagnoses, a chi-square was conducted on the PMR and SMR groups to assess differences in the number of autism diagnoses between groups while controlling for variation due to chance (Hinkle, Wiersma, & Jurs, 1998). The chi-square indicated a significant relationship between level of mental retardation and number of DSM-IV diagnoses, $\chi^2(1) = 13.04$, $p < .001$. As can be viewed in Table 7, 19 individuals with profound mental retardation were labeled autistic by the DSM-IV, as opposed to only 4 individuals with severe mental retardation.

Table 7

Chi-Square for Level of MR and DSM-IV Checklist Diagnostic

Level of MR	DSM-IV	
	Autism	No Autism
Profound	19	27
Severe	4	42

Discussion

The current study focused on three objectives. The first goal was to determine whether the CARS and DSM-IV Checklist for autism could be used reliably with individuals with severe and profound mental retardation. In addition, this objective was designed to assess whether these two instruments are appropriate when direct-care staff are used as informants. The second goal of this study was to determine the concordance between the CARS and DSM-IV autism “diagnoses.” Because these two instruments measure the same construct, we expected their “diagnoses” of people with mental retardation to agree in large measure. Finally, the third goal of the present study was to compare individuals with profound mental retardation (PMR) to those with severe mental retardation (SMR) as a preliminary examination of the appropriateness of the use of the CARS and DSM-IV on these populations when autism was being assessed. These particular goals were chosen as a step towards quantifying the reliability and concordance of these scales in this population.

The first objective of this study, assessing the reliability of the CARS and DSM-IV Checklist, entailed the analysis of two types of reliability for each instrument: test-retest and interrater. In order for a scale to be clinically valuable it must provide consistent ratings (Groth-Marnat, 1999). Test-retest reliability was conducted using direct-care staff ratings. The test-retest reliability of the CARS was found to be .86, which, according to Cicchetti (1994), is an excellent level of clinical significance. In addition, the test-retest reliability of the individual CARS items ranged from .42 for Activity Level to .85 for Verbal Communication, and all correlations coefficients had fair to excellent clinical significance (Cicchetti, 1994). Strong item test-retest lends itself to a stronger overall test-retest reliability of the instrument.

A high correlation coefficient in this case indicates that the first and second ratings of the direct-care staff were stable over time. The test-retest reliability found in the current study is comparable to the reliability of .88 reported by Schopler et al. (1980) in the initial CARS psychometrics performed on children. Good test-retest reliability is a necessary component of the psychometrics of any instrument, so this result attests to the stability of direct-care staff reports when administering the CARS on behalf of individuals with mental retardation.

The test-retest reliability of the DSM-IV was lower than that found for the CARS. The resulting correlation, $\kappa = .64$, is considered to have good clinical significance (Cicchetti, 1994), but is still rather low for clinical use. Individual item kappa correlations ranged from poor for Lack of Make-Believe play ($\kappa = .15$) to excellent for Stereotyped Motor Mannerisms ($\kappa = .82$). Items that were easily observable, such as stereotyped behavior and delay in or lack of spoken language were the most stable upon retest. The weakest correlation obtained was for an item that was more subjective in nature, Lack of Make-Believe Play.

The source of low reliability in this case is most likely attributable more to the readability of the DSM-IV Checklist rather than the DSM-IV items' meaning. The DSM-IV was designed for use by trained clinicians with advanced college degrees (APA, 1994) and this is reflected in the language used in describing the criteria. The DSM-IV criteria utilize some professional jargon, (such as echolalia, stereotyped patterns of interest, and social imitative play,) as well as double negatives and other potentially confusing phrasing. Residential facilities' direct-care staff members come from a variety of educational backgrounds (Adkins, Singh, McKeegan, Lanier, & Oswald, 2002), with a potential for vastly different experiences. As a result, the DSM-IV may have been difficult to understand for some of the direct-care staff. Support for this argument can be drawn from observing the test-retest reliability of the individual DSM-IV Checklist items.

The two lowest correlations were for the items dealing with spontaneous make-believe and social imitative play and stereotyped and restricted patterns of interest. Whereas, Stereotyped Motor Mannerisms, a behavior that is easily observable and more familiar to the direct-care staff, received the strongest test-retest correlation coefficient. Controlling for the wording of criterion was beyond the scope of this study, but may be an area for research attention in the future.

The second type of reliability studied was interrater. For paired direct-care staff, the interrater reliability of the CARS ($r = .42$) and DSM-IV diagnoses ($\kappa = .32$), while both significant at the $p < .01$ level, were both poor correlations (Cicchetti, 1994). For the CARS, the only good coefficient was for Verbal Communication. The remaining items could all be classified as fair for Relating to People, Object Use, and Activity Level and poor for the remaining items (Cicchetti, 1994). Similar results were found for the items on the DSM-IV. The only good correlation was for Delay in the Development of Spoken Language. All other correlations were poor.

These results question the appropriateness of the CARS and DSM-IV when direct-care staff is used as raters. To compile more information to answer this question, direct-care staff ratings were compared to those made by trained, masters-level mental health professionals for each of the participants. This comparison yielded similar results. The interrater reliability for paired direct-care staff and masters-level mental health professionals was low for both the CARS ($r = .50$) and DSM-IV ($\kappa = .31$). Individual CARS items generally had lower correlation coefficients for direct-care/professional pairs than for direct-care staff pairs. The only excellent correlation was found for Verbal Communication. All other correlation coefficients were poor (Cicchetti, 1994). Correlation coefficients for DSM-IV items could also all be categorized as

poor, except for Delay in the Development of Spoken Language, which was of good clinical significance (Cicchetti, 1994).

The data suggest that direct-care staff generally do not agree with each other on CARS and DSM-IV ratings, and also do not agree with masters-level mental health professionals. The source of this lack of concordance is unclear at this time. It is possible that there is something inherent in these instruments that makes agreement difficult, such as the inclusion of items unrelated to the construct being studied, the instability of the behaviors over time or across situations. The low interrater agreement between direct-care staff can be seen as support for this explanation.

On the other hand, the lack of agreement may be due to the differing education and experience between the groups of raters. Support for this second argument can be found in the literature. Volkmar et al. (1994) performed a study on DSM-IV interrater reliability during the field trials for autism in the DSM-IV. Two experienced clinicians' ratings were compared on the DSM-III, DSM-III-R, and ICD-10 and found that the DSM-III had the lowest interrater reliability ($\kappa = .59$), followed by the DSM-III-R ($\kappa = .67$) and the ICD-10 ($\kappa = .68$). The ability of the clinicians to reliably differentially diagnosis autism from other PDDs was very high ($\kappa = .95$). However, when an experienced rater was compared to an inexperienced rater, the resulting correlation was only fair ($\kappa = .59$). This line of research was continued by Klin et al. (2000) who compared experienced raters, (defined as clinicians who had been involved in the assessment and diagnosis of over 25 individuals,) with inexperienced raters, (nurses, social workers, special educators, etc.) Klin et al. (2000) found that inexperienced raters had markedly lower reliability coefficients than those with experience when using the DSM-IV.

By comparing direct-care staff with professionals, our data extends the conclusion that experienced and inexperienced raters do not agree to DSM-IV to a more applicable group of informants in residential settings. These results provide additional support for the possibility that these two instruments are not appropriate for use with direct-care staff. The studies by Volkmar et al. (1994) and Klin et al. (2000), coupled with poor direct-care staff interrater data reported here, is in line with the current study's findings on lack of agreement between raters with varying levels of experience and education. Obviously, this presents a practical concern, since these same staff shortcomings are likely to be common throughout the field. Since it was not possible to obtain multiple masters-level informants for the subjects in this study, it is not known whether the interrater agreement between informants with similar education and experience would be greater.

For the CARS, the only interrater reliability data published was in reference to trained observers. Schopler et al. (1980) compared the individual item scores from two independent observers and found an average interrater reliability of .71. Although the CARS has been reported to be valid when used by other raters (Schopler et al., 1988), no reliability data is available at this time. These results conclude that the CARS has poor interrater reliability when direct-care staff and professionals are compared.

The current study has provided information not found elsewhere in the literature. The reliability portion of this study provided information that heretofore was unavailable regarding the DSM-IV and CARS. Excellent test-retest reliability of the CARS and fair test-retest reliability of the DSM-IV Checklist using direct-care staff was found. However, test-retest data did not paint the full picture since interrater reliability may imply that the source of disagreement is test error or appropriateness of the informants (Groth-Marnat, 1999). Unfortunately, in many

applied settings there are simply not sufficient numbers of professionals with the requisite training to do these evaluations. Direct-care staff is often relied upon for information due to their daily interactions with the clients and the limited amount of time psychologists have for each client in their caseload. Previous research has brought doubt upon the accuracy and agreement in direct-care staff ratings in numerous areas, such as challenging behaviors and reinforcer preferences (Green et al., 1988; McGill et al., 2001; Parsons & Reid, 1990). The results of this study also put into question the appropriateness of direct-care staff ratings of autism. Future research into the reliability of the CARS and DSM-IV in this population should attempt to replicate the current findings, with an emphasis on the reliability of masters-level or higher trained professionals.

The second objective of the present study was to examine the concordance between “diagnoses” made by the CARS and DSM-IV Checklist. It is important to note at this time that what follows is a summary of the data collected. Due to the lack of reliability found in the previous portion of the study, validity cannot be considered at this time.

Since these two instruments purport to measure the same construct, one would expect high levels of diagnostic agreement (Groth-Marnat, 1999). Diagnostic concordance was measured for both direct-care staff and professional ratings. For both types of raters, concordance between CARS and DSM-IV diagnoses of “autism” and “no autism” was significant. This finding indicates that the CARS and DSM-IV Checklist agreed on diagnoses at a rate greater than that expected by chance. It appears as though these two instruments are measuring the same construct in this population of individuals with mental retardation. Other research has shown that the CARS concurs with the DSM-III-R. In a study by Sevin et al. (1991), CARS diagnoses (i.e. using the cutoff of 30) agreed 92% of the time with DSM-III-R

criterion-based diagnoses. In a similar study, Van Bourgondien, Marcus and Schopler (1992) found a significant relationship between DSM-III-R and CARS diagnoses made by clinicians. As clinicians attempt to corroborate their diagnoses by using multiple measures (Haynes & O'Brien, 2000), it is vital that research is available on the nature of their concordance.

It is important to note at this time that the question being answered here is not that of validity. That is, we have no direct evidence to support that this construct is autism. With the current state of the literature and empirical evidence on adults with autism, it is not possible to consider either of these instruments the criterion for tests of validity. The objective of this portion of the study was to determine concordance between the two instruments. The result of this analysis was that the CARS and DSM-IV agree on diagnosis the majority of the time. With knowledge of concordance, attention should now be turned to evaluating the validity of these two instruments.

The final objective of the present study was to compare the PMR and SMR groups on total CARS scores and DSM-IV diagnoses. Part of determining the utility of the CARS and DSM-IV in this population is establishing whether these instruments can differentiate between characteristics of mental retardation and autism. An independent samples t-test found a significant difference between total CARS scores for the PMR and SMR groups. Specifically, the average CARS score for the PMR group ($M = 31.97$) was higher than that for the SMR group ($M = 22.44$). A similar result was found for DSM-IV diagnoses. A chi-square analysis found a significant relationship between group membership and diagnostic category. As can be seen in Table 7, 19 of the individuals in the PMR group were diagnosed with "autism," as opposed to only 4 individuals in the SMR group. Thus, it appears as though individuals with profound

mental retardation receive higher CARS scores, the average of which is slightly above the cutoff of 30, as well as significantly more DSM-IV autism diagnoses.

A possible explanation for these findings is that a greater degree of symptom overlap occurs between autism and profound mental retardation than between autism and severe mental retardation. Many items found in the CARS and DSM-IV Checklist are similar to behaviors found in mental retardation. However, it appears as though both instruments were capable of differentiating between mental retardation and autism in the SMR group, as evinced by the mean CARS score falling well below the cutoff of 30 and only 4 individuals being placed in the autism group on the DSM-IV. This was not the case in the PMR group, as the mean CARS score was slightly above the cutoff and 19 out of 46 of the participants were labeled autistic by the DSM-IV. Individuals with profound mental retardation are more likely to have deficits in communication and social skills, as well as being more likely to exhibit stereotyped and other problem behaviors (APA, 1994). All of these areas are also considered to be characteristic of autism on the CARS and DSM-IV. Further research is needed to determine whether it is possible to differentiate autism from profound mental retardation in adults.

The results of the current study extend previous research on the overlap of the behavioral characteristics of autism and mental retardation. In Lund's (1986) sample of 302 adults with mental retardation, 27% exhibited social withdrawal, 28% had abnormal language, and 22% had stereotyped behavior. However, only 7.6% met the criteria for autism. This study did not report the diagnostic criteria used to diagnose autism, so comparison to the current results is difficult. In another study, Janicki and Jacobson (1983), using a sample of 314 adults with autism, found that 43% also had profound mental retardation, 30% severe, 17% moderate, and 10% mild. Similar to Lund (1986), Janicki and Jacobson (1983) failed to report diagnostic criteria used to

diagnosis autism in their sample. The present study was the first to examine overlap using commonly used diagnostic instruments. Future research using this sample should investigate the nature of the overlap found in both the current and previous studies.

Researchers have long implicitly treated autism as a disorder specific to childhood. As a result, clinicians currently have no empirically supported instruments to aid in their evaluations of adults. Given the increased life expectancy and enhanced community integration of this population, assessment data such as this becomes of far greater importance. The current study established the reliability of two commonly used instruments, the CARS and DSM-IV, on individuals with mental retardation using multiple informants, in addition to determining the concordance between the DSM-IV and CARS and providing a preliminary analysis of the appropriateness of these two instruments on individuals with severe and profound mental retardation. Researchers should continue to focus on adults with mental retardation who may not have received dual diagnoses as children and as a means to assess shifting symptom patterns. Our ability to serve this population will improve by continuing to study the psychometrics of diagnostic instruments and the overlap between symptoms of mental retardation and psychopathology.

References

- Abt, I. (1965). History of pediatrics. Philadelphia: W.B. Saunders.
- Achenbach, T.M., McConaughy, S.H., & Howell, C.T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. Psychological Bulletin, 101, 213-232.
- Adkins, A.D., Singh, N.N., McKeegan, G.F., Lanier, A.E., & Oswald, D.P. (2002). Behavior treatment programs, readability, and treatment outcomes. Research in Developmental Disabilities, 23, 253-265.
- American Psychiatric Association. (1980). Diagnostic and statistical manual of mental disorders (3rd ed.). Washington, D.C.: Author.
- American Psychiatric Association. (1987). Diagnostic and statistical manual of mental disorders (3rd ed., revised). Washington, D.C.: Author.
- American Psychiatric Association. (1994). Diagnostic and statistical manual of mental disorders (4th ed.). Washington, D.C.: Author.
- Anderson, G.M., & Hoshino, Y. (1987). Neurochemical studies of autism. In D.J. Cohen A.M. Donnellan, & R. Paul (Eds.), Handbook of autism and pervasive developmental disorders (pp.166-191). New York: Wiley.
- Baird, P.A., & Sadovnick, A.D. (1985). Mental retardation in over half a million consecutive live births—An epidemiological study. American Journal of Mental Deficiency, 89, 323-330.
- Bartak, L., & Rutter, M. (1974). Differences between mentally retarded and normally intelligent autistic children. Journal of Autism and Childhood Schizophrenia, 6, 109-120.
- Bartak, L., Rutter, M., & Cox, A. (1977). A comparative study of infantile autism and specific developmental receptive language disorders: III. Discriminate function analysis. Journal of Autism and Childhood Schizophrenia, 7, 383-396.
- Bernsen, A.H. (1976). Severe mental retardation in the county of Arhus, Denmark. A community study on prevalence and provision of service. Acta Psychiatrica Scandinavica, 54, 43-66.
- Bettelheim, B. (1967). The empty fortress. New York: Free Press.
- Borthwick-Duffy, S. (1994). Epidemiology and prevalence of psychopathology in people with mental retardation. Journal of Consulting and Clinical Psychology, 62, 17-27.

- Borthwick-Duffy, S.A., & Eyman, R.K. (1990). Who are the dually diagnosed? American Journal of Mental Retardation, 94, 586-595.
- Brooker, A.E., & Mareth, T.R. (1982). Infantile autism: Clinical features, diagnosis, etiology, and prognosis—a research review. Psychological Reports, 50, 587-592.
- Cicchetti, D.V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychological Assessment, 6, 284-290.
- Cohen, D.J., Caparulo, B.K., Gold, J.R., Waldo, M.C., Shaywitz, B.A., Ruttenberg, B.A., & Rimland, B. (1978). Agreement in diagnosis: Clinical assessment and behavior rating scales for pervasively disturbed children. Journal of the American Academy of Child Psychiatry, 17, 589-603.
- Cohen, D.J., Paul, R., & Volkmar, F.R. (1986). Issues in the classification of pervasive and other developmental disorders: Toward DSM-IV. Journal of the American Academy of Child Psychiatry, 25, 213-220.
- Creak, M. (1961). Schizophrenia syndrome in childhood: Progress report of a working party. Cerebral Palsy Bulletin, 3, 501-504.
- Creak, M. (1964). Schizophrenic syndrome in childhood: Further progress report of a working party (April, 1964). Developmental Medicine and Child Neurology, 6, 530-535.
- Curcio, F. (1978). Sensorimotor functioning and communication in mute autistic children. Journal of Autism and Childhood Schizophrenia, 3, 281-292.
- Darragh, P.M. (1982). The prevalence and prevention of severe mental handicap in Northern Ireland. Irish Medical Journal, 75, 147-155.
- Dawson, J.E., Matson, J.L., & Cherry, K.E. (1998). An analysis of maladaptive behaviors in persons with autism, PDD-NOS, and mental retardation. Research in Developmental Disabilities, 19, 439-448.
- DeMyer, M.K. (1979). Parents and children in autism. Washington, D.C.: Victor H. Winston & Sons.
- DeMyer, M.K., Churchill, D.W., Pontius, W., & Gilkey, K.M. (1971). A comparison of five diagnostic systems for childhood schizophrenia and infantile autism. Journal of Autism and Childhood Schizophrenia, 1, 175-189.
- Elkin, E. (1977). Historical perspectives. In National Association for Retarded Citizens (Ed.), A national forum on residential services (pp. 5-12). Arlington, TX: Author.

- Factor, D.C., Freeman, N.L., & Kardish, A. (1989). A comparison of DSM-II and DSM-III-R criteria for autism. Journal of Autism and Developmental Disorders, 19, 637-640.
- Fee, V.E., & Matson, J.L. (1992). Definition, classification, and taxonomy. In J.K. Luiselli, J.L. Matson, & N.N. Singh (Eds.), Self-injurious behavior: Analysis, assessment, and treatment (pp.3-20). New York: Springer-Verlag.
- Fiedler, L. (1978). Freaks. New York: Simon and Schuster.
- Fish, B. (1976). Biological disorders in infants at risk for schizophrenia. In E.R. Ritvo (Ed.), Autism: Diagnosis, current research and management (pp. 111-146). New York: Spectrum.
- Fishbach, M., & Hull, J.T. (1982). Mental retardation in the province of Manitoba. Canada's Mental Health, 30, 16-19.
- Folstein, S., & Rutter, M. (1977). Infantile autism: A genetic study of 21 twin pairs. Journal of Child Psychology and Psychiatry, 18, 297-321.
- Freeman, B.J., Ritvo, E.R., & Schroth, P.C. (1984). Behavior assessment of the syndrome of autism: Behavioral Observation System. Journal of the American Academy of Child Psychiatry, 23, 588-594.
- Gardner, W.I., & Cole, C.L. (1990). Aggression and related conduct difficulties. In J.L. Matson (Ed.), Handbook of behavior modification with the mentally retarded (2nd ed) (pp.225-248). New York: Plenum Press.
- Gillberg, C. (1984). Infantile autism and other childhood psychoses in a Swedish region: Epidemiological aspects. Journal of Child Psychology and Psychiatry, 25, 35-43.
- Green, C.W., Reid, D.H., White, L.K., Halford, R.C., Brittain, D.P., & Gardner, S.M. (1988). Identifying reinforcers for persons with profound handicaps: Staff opinion versus systematic assessment of preferences. Journal of Applied Behavior Analysis, 21, 31-43.
- Groth-Marnat, G. (1999). Handbook of psychological assessment. New York: John Wiley & Sons, Inc.
- Gustavson, K.H., Hagberg, B., Hagberg, G., & Sars, K. (1977). Severe mental retardation in a Swedish county. II. Etiological and pathogenic aspects of children born 1959-70. Neuropaediatrie, 8, 293-304.
- Hayes, H. (1969). Smiling through the apocalypse. New York: McCall.
- Haynes, S.N., & O'Brien, W.H. (2000). Principles and practice of behavioral assessment. New York: Kluwer Academic/Plenum Publishers.

- Heaton-Ward, A. (1977). Psychosis in mental handicap. British Journal of Psychiatry, 130, 525-533.
- Hermelin, J.N., & O'Conner, N. (1970). Psychological experiments with autistic children. Oxford: Pergamon Press.
- Hertzig, M.E., Snow, M.E., New, E., & Shapiro, T. (1990). DSM-III and DSM-III-R diagnosis of autism and pervasive developmental disorder in nursery school children. Journal of the American Academy of Child and Adolescent Psychiatry, 29, 123-126.
- Hill, B.K., & Bruninks, R.H. (1984). Maladaptive behavior of mentally retarded individuals in residential facilities. American Journal of Mental Deficiency, 88, 380-387.
- Hinkle, D.E., Wiersma, W., & Jurs, S.G. (1998). Applied statistics for the behavioral sciences. New York: Houghton Mifflin Company.
- Hitzig, W. (1987). Community living alternatives for persons with autism and related severe behavioral problems. In D.J. Cohen A.M. Donnellan, & R. Paul (Eds.), Handbook of autism and pervasive developmental disorders (pp. 396-417). New York: Wiley.
- Hobson, R.P. (1989). Beyond cognition: A theory of autism. In G. Dawson (Ed.), Autism: Nature, diagnosis, and treatment (pp.22-48). New York: Guilford.
- Holmes, N., Shah, A., & Wing, L. (1982). The Disability Assessment Schedule: A brief screening device for use with the mentally retarded. Psychological Medicine, 12, 879-890.
- Howlin, P., & Rutter, M. (1987). Treatment of autistic children. London: Wiley.
- Jacobson, J.W., & Ackerman, L.J. (1990). Differences in adaptive functioning among people with autism or mental retardation. Journal of Autism and Developmental Disorders, 20, 205-219.
- Janicki, M.P., & Jacobson, J.W. (1983). Selected clinical features and service characteristics of autistic adults. Psychological Reports, 52, 387-390.
- Kanner, L. (1943). Autistic disturbances of affective contact. Nervous Child, 2, 217-250.
- Kanner, L. (1951). The conception of wholes and parts in early infantile autism. American Journal of Psychiatry, 108, 23-26.
- Kanner, L. (1971). Follow-up of 11 autistic children originally seen in 1943. Journal of Autism and Childhood Schizophrenia, 1, 119-145.

- Klin, A., Lang, J., Cicchetti, D.V., & Volkmar, F.R. (2000). Brief report: Interrater reliability of clinical diagnosis and DSM-IV criteria for autistic disorder: Results of the DSM-IV field trial. *Journal of Autism and Developmental Disorders*, 30, 163-167.
- Krug, D.A., Arick, J., & Almond, P. (1980). Behavior checklist for identifying severely handicapped individuals with high levels of autistic behavior. *Journal of Child Psychology and Psychiatry*, 21, 221-229.
- Larson, S.A., Lakin, K.C., Anderson, L., Kwak, N., Lee, J.H., & Anderson, D. (2001). Prevalence of mental retardation and developmental disabilities: Estimates from the 1994/1995 National Health Interview Survey Disability Supplements. *American Journal on Mental Retardation*, 106, 231-252.
- Laxova, R., Ridler, M.A.C., & Bowen-Bravery, M. (1977). An etiological survey of the severely retarded Hertfordshire children who were born between January 1, 1965 and December 31, 1967. *American Journal of Medical Genetics*, 1, 75-86.
- Lindsey, M.P., & Russell, C.M. (1981). Mental handicap in the county of Cornwall: Prevalence and the use of services. *Journal of Mental Deficiency Research*, 25, 77-87.
- Leakey, R., & Lewin, R. (1978). *People of the lake*. New York: Doubleday.
- Lotter, V. (1966). Epidemiology of autistic conditions in young children: I. Prevalence. *Social Psychiatry*, 1, 124-137.
- Lotter, V. (1967). Epidemiology of autistic conditions in young children: II. Some characteristics of the parents and children. *Social Psychiatry*, 1, 163-173.
- Lotter, V. (1974). Social adjustment and placement of autistic children in Middlesex: A follow-up. *Journal of Autism and Childhood Schizophrenia*, 4, 11-32.
- Lund, J. (1986). Behavioral symptoms and autistic psychosis in the mentally retarded adult. *Acta Psychiatria Scandinavia*, 73, 420-428.
- Mahler, M.S. (1952). On childhood psychosis and schizophrenia: Autistic and symbiotic psychoses. *The Psychoanalytic Study of the Child*, 1, 286-305. (Reprinted in S.I. Harrison & J. F. McDermott (Eds.), *Childhood psychopathology: An anthology of basic reading*. New York: International Universities Press.)
- Masters, J.C., & Miller, D.E. (1970). Early infantile autism: A methodological critique. *Journal of Autism and Childhood Schizophrenia*, 8, 162-167.
- Matson, J.L. (1985). Biosocial theory of psychopathology: A three by three factor model. *Applied Research in Mental Retardation*, 6, 199-227.

- Matson, J.L. (1995a). Matson Evaluation of Social Skills in Individuals with Severe Retardation. Baton Rouge: Scientific Publishers, Inc.
- Matson, J.L. (1995b). Diagnostic Assessment for the Severely Handicapped-II. Baton Rouge, LA: Scientific.
- Matson, J.L., & Bamburg, J.W. (1998). Reliability of the Assessment of Dual Diagnosis (ADD). Research in Developmental Disabilities, 19, 89-95.
- Matson, J.L., & Frame, C.L. (1986). Psychopathology among mentally retarded children and adolescents. Beverly Hills, CA: Sage.
- Matson, J.L., Gardner, W.I., Coe, D.A., & Sovner, R. (1991). A scale for evaluating emotional disorders in severely and profoundly mentally retarded persons. British Journal of Psychiatry, 159, 404-409.
- Matson, J.L., Sevin, J.A., Fridley, D., & Love, S.R. (1990). Increasing spontaneous language in three autistic children. Journal of Applied Behavior Analysis, 23, 227-234.
- Matson, J.L., Smiroldo, B.B., & Hastings, T.L. (1998). Validity of the autism/pervasive developmental disorder subscale of the Diagnostic Assessment for the Severely Handicapped-II. Journal of Autism and Developmental Disorders, 28, 77-81.
- McGill, P., Hughes, D., Teer, K., & Rye, L. Variability in staff reports of the frequency of challenging behavior. Research in Developmental Disabilities, 22, 221-231.
- McLaren, J., & Bryson, S.E. (1987). Review of recent epidemiological studies of mental retardation: Prevalence, associated disorders, and etiology. American Journal of Mental Retardation, 92, 243-254.
- McQueen, P.C., Spence, M.W., Garner, J.B., Pereira, L., & Winsor, E.J. (1987). Prevalence of major mental retardation and associated disabilities in the Canadian Maritime Provinces. American Journal of Mental Deficiency, 91, 460-466.
- Menolascino, F.J. (1977). Challenged in mental retardation: Progressive ideology and services. New York: Human Services Press.
- Menolascino, F.J. (1988). Mental illness in the mentally retarded: Diagnostic and treatment issues. In J.A. Stark, F.J. Menolascino, M.H. Albarelli, & V.C. Gray (Eds.), Mental retardation and mental health (pp. 109-123). New York: Springer-Verlag.
- Mesibov, G.B., Schopler, E., Schaffer, B., & Michal, N. (1989). Use of the Childhood Autism Rating Scale with autistic adolescents and adults. Journal of the American Academy of Child Adolescent Psychiatry, 28, 538-541.

Morgan, S. (1988). Diagnostic assessment of autism: A review of objective scales. Journal of Psycheducational Assessment, 6, 130-151.

Mundy, P., & Sigman, M. (1989). Specifying the nature of the social impairment in autism. In G. Dawson (Ed.), Autism: Nature, diagnosis, and treatment (pp.3-21). New York: Guilford.

Nihira, K., Foster, R., Shellhaas, M., & Leland, H. (1969). AAMD Adaptive Behavior Scale. Washington, D.C.: American Association on Mental Deficiency.

Njardvik, U., Matson, J.L., & Cherry, K.E. (1999). A comparison of social skills in adults with autistic disorder, pervasive developmental disorder not otherwise specified, and mental retardation. Journal of Autism and Developmental Disorders, 29, 287-295.

Ornitz, E.M., Guthrie, D., & Farley, A.J. (1978). The early symptoms of childhood autism. In G. Serban (Ed.), Cognitive defects in the development of mental illness. New York: Brunner/Mazel.

Parks, S.L. (1988). Psychometric instruments available for the assessment of autism. In E. Schopler & G.B. Mesibov (Eds.), Diagnosis and assessment in autism (pp. 123-136). New York: Plenum Press.

Parsons, M.B., & Reid, D.H. (1990). Assessing food preferences among persons with profound mental retardation: Providing opportunities to make choices. Journal of Applied Behavior Analysis, 23, 183-195.

Pauls, D.L. (1987). The familiarity of autism and related disorders: A review of the evidence. In D.J. Cohen, A.M. Donnellan, & R. Paul (Eds.), Handbook of autism and developmental disorders (pp. 192-207). New York: John Wiley & Sons.

Phelps, L., & Grabowski, J. (1991). Autism: Etiology, differential diagnosis, and behavioral assessment update. Journal of Psychopathology and Behavioral Assessment, 13, 107-125.

Powers, M.D. (1988). Behavioral assessment of autism. In E. Schopler & G.B. Mesibov (Eds.), Diagnosis and assessment in autism (pp. 139-165). New York: Plenum Press.

Powers, M.D., & Handleman, J.S. (1984). Behavioral assessment of severe developmental disabilities. Rockville, MD: Aspen Systems Corp.

Reiss, S. (1993). Assessment of psychopathology in persons with mental retardation. In J.L. Matson & R.P. Barrett (Eds.), Psychopathology in the mentally retarded, 2nd Edition (pp.17-39). Needham Heights, MA: Allyn & Bacon.

Rimland, B. (1964). Infantile autism: The syndrome and its implications for a neural theory of behavior. New York: Appleton-Century-Crofts.

- Rimland, B. (1968). On the objective diagnosis of infantile autism. Acta Paedopsychiatrica, 35, 146-161.
- Rimland, B. (1971). The differentiation of childhood psychoses: An analysis of checklist for 2,218 psychotic children. Journal of Autism and Childhood Schizophrenia, 8, 162-167.
- Ritvo, E.R., & Freeman, B.J. (1977). National Society for Autistic Children definition of the syndrome of autism. Journal of Pediatric Psychology, 24, 142-145.
- Ritvo, E.R., Freeman, B.J., Mason-Brothers, A., Mo, A., & Ritvo, A.M. (1985). Evidence for autosomal recessive inheritance in 46 families with multiple incidence of autism. American Journal of Psychiatry, 142, 187-192.
- Rodrigue, J.R, Morgan, S.M., & Geffken, G.R. (1991). A comparative evaluation of adaptive behavior in children and adolescents with autism, Down syndrome, and normal development. Journal of Autism and Developmental Disorders, 21, 187-196.
- Rojahn, J., & Sisson, L.A. (1990). Stereotyped behavior. In J.L. Matson (Ed.), Handbook of behavior modification with the mentally retarded (2nd ed) (pp.181-213). New York: Plenum Press.
- Rumsey, J.M., Rapaport, J.L., & Sceery, W.R. (1985). Autistic children as adults: Psychiatric, social and behavioral outcomes. Journal of American Academy of Child and Adolescent Psychiatry, 24, 465-473.
- Ruttenberg, B.A., Dratman, M.L., Frankno, J., & Wenar, C. (1966). An instrument for evaluating autistic children. Journal of the American Academy of Child Psychiatry, 5, 453-478.
- Ruttenberg, B.A., Kalish, B.I., Wenar, C., & Wolf, E.G. (1977). Behavior rating instrument for autistic and other atypical children. (Rev. ed.). Philadelphia: Developmental Center for Autistic Children.
- Rutter, M. (1978). Diagnosis and definition of childhood autism. Journal of Autism and Childhood Schizophrenia, 8, 139-161.
- Rutter, M. (1989). Isle of Wight revisited: Twenty-five years of child psychiatric Epidemiology. Journal of the American Academy of Child and Adolescent Psychiatry, 28, 633-653.
- Rutter, M., & Garmezy, N. (1983). Developmental psychopathology. In E.M. Hetherington (Ed.), Handbook of child psychology, Vol. 4. New York: Wiley.
- Rutter, M., Greenfield, D., & Lockyer, L. (1967). A five to fifteen year follow-up of infantile psychosis: II. Social and behavioral outcome. British Journal of Psychiatry, 113, 1183-1199.

Rutter, M., Graham, P., & Yule, W. (1970). A neuropsychiatric study in childhood. London: Heinemann.

Rutter, M., Tizard, J., & Whitmore, K. (1970). Education, health, and behavior. London: Longmans.

Scheerenberger, R.C. (1983). A history of mental retardation. Baltimore: Brookes.

Schopler, E., Reichler, R.J., DeVellis, R.F., & Daly, K. (1980). Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS). Journal of Autism and Developmental Disorders, 10, 91-103.

Schopler, E., Reichler, R.J., & Renner, B.R. (1988). The Childhood Autism Rating Scale (CARS). Los Angeles: Western Psychological Services.

Sevin, J.A., Matson, J.L., Coe, D.A., Fee, V.E., & Sevin, B.M. (1991). A comparison and evaluation of three commonly used autism scales. Journal of Autism and Developmental Disorders, 21, 417-432.

Shah, A., Holmes, N., & Wing, L. (1982). Prevalence of autism and related conditions in adults in a mental handicap hospital. Applied Research in Mental Retardation, 3, 303-317.

Sparrow, S., Balla, D., & Cicchetti, D. (1984). Vineland Adaptive Behavior Scales. Circle Pines, MN: American Guidance Service.

Stone, W.L., & Caro-Martinez, L.M. (1990). Naturalistic observations of communication in autistic children. Journal of Autism and Developmental Disorders, 20, 437-454.

Sturmey, P., & Sevin, J.A. (1994). Defining and assessing autism. In J.L. Matson (Ed.), Autism in children and adults: Etiology, assessment, and intervention (pp. 13-36). Pacific Grove, CA: Brooks/Cole.

Teal, M.B., & Wiebe, M.J. (1986). A validity analysis of selected instruments used to assess autism. Journal of Autism and Developmental Disorders, 16, 485-494.

Thomas, J.R. (1994). Quality care for individuals with dual diagnosis: The legal and ethical imperative to provide qualified staff. Mental Retardation, 32, 356-361.

Treffert, D. (1970). Epidemiology of infantile autism. Archives of General Psychiatry, 22, 431-438.

Ungerer, J.A. (1989). The early development of autistic children: Implications for defining primary deficits. In G. Dawson (Ed.), Autism: Nature, diagnosis, and treatment (pp. 75-91). New York: Guilford.

- Van Bourgondien, M.E., Marcus, L.J., & Schopler, E. (1992). Comparison of DSM-III-R and Childhood Autism Rating Scale diagnoses of autism. *Journal of Autism and Developmental Disorders*, 22, 493-506.
- Volkmar, F.R. (1987). Social development. In D.J. Cohen, A.M. Donnellan, & R. Paul (Eds.), *Handbook of autism and developmental disorders* (pp. 41-60). New York: John Wiley & Sons.
- Volkmar, F.R., & Anderson, G.M. (1989). Neurochemical perspectives on infantile autism. In G. Dawson (Ed.), *Autism: Nature, diagnosis, and treatment* (pp.208-224). New York: Guilford Press.
- Volkmar, F.R., Bregman, J., Cohen, D.J., & Cicchetti, D.V. (1988). DSM-III and DSM-III-R diagnoses of autism. *American Journal of Psychiatry*, 145, 1404-1408.
- Volkmar, F.R., Cicchetti, D.V., Bregman, J., & Cohen, D.J. (1992a). Three diagnostic systems for autism: DSM-III, DSM-III-R, and ICD-10. *Journal of Autism and Developmental Disorders*, 22, 483-492.
- Volkmar, F.R., Cicchetti, D.V., Bregman, J., & Cohen, D.J. (1992b). Developmental aspects of DSM-III-R criteria for autism. *Journal of Autism and Developmental Disorders*, 22, 657-662.
- Volkmar, F.R., Cicchetti, D.V., Dykens, E., Sparrow, S.S., Leckman, J.F., & Cohen, D.J. (1988). An evaluation of the Autism Behavior Checklist. *Journal of Autism and Developmental Disorders*, 18, 81-97.
- Volkmar, F.R., Klin, A., Siegel, B., Szatmari, P., Lord, C., Campbell, M., Freeman, B.J., Cicchetti, D.V., Rutter, M., Kline, W., Buitelaar, J., Hattab, Y., Fombonne, E., Fuentes, J., Werry, J., Stone, W., Kerbeshian, J., Hoshino, Y., Bregman, J., Loveland, K., Szymanski, L., Towbin, K. (1994). Field trial for autistic disorder in DSM-IV. *American Journal of Psychiatry*, 151, 1361-1367.
- Volkmar, F.R., & Schwab-Stone, M. (1996). Annotation: Childhood disorders in DSM-IV. *Journal of Child Psychology and Psychiatry*, 37, 779-784.
- Wing, L. (1981). Sex ratios in early childhood autism and related conditions. *Psychiatry Research*, 5, 129-137.
- Wing, L., & Gould, J. (1979). Severe impairments of social interaction and associated abnormalities in children: Epidemiology and classification. *Journal of Autism and Developmental Disorders*, 9, 11-29.
- Wolf, E.G., Wenar, C., & Ruttenberg, B.A. (1972). A comparison of personality variables in autistic and mentally retarded children. *Journal of Autism and Childhood Schizophrenia*, 2, 92-108.

World Health Organization. (1996). Multiaxial classification of child and adolescent psychiatric disorders: The ICD-10 classification of mental and behavioral disorders in children and adolescents. Geneva: Author.

Wright, E.C. (1982). The presentation of mental illness in mentally retarded adults. British Journal of Psychiatry, 141, 496-502.

Yuwiler, A., Geller, E., & Ritvo, E. (1985). Biochemical studies of autism. In E. Lajtha (Ed.), Handbook of neurochemistry (pp. 671-691). New York: Plenum.

Appendix A

DSM-IV Checklist

1 = YES 0 = NO

- _____ 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 relations 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)
- _____ 4. Lack of social or emotional reciprocity (give-and-take)
- _____ 5. 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)
- _____ 6. In individuals with adequate speech, marked impairment in the ability to initiate or sustain a conversation with others
- _____ 7. Stereotyped and repetitive use of language
- _____ 8. Lack of varied, spontaneous make-believe play or social imitative play appropriate to developmental level
- _____ 9. Encompassing preoccupation with one or more stereotyped and restricted patterns of interest that is abnormal either in intensity or focus
- _____ 10. Apparently inflexible adherence to specific, nonfunctional routines or rituals
- _____ 11. Stereotyped and repetitive motor mannerism (e.g., hand or finger flapping or twisting, or complex whole-body movements)
- _____ 12. Persistent preoccupation with parts of objects

Appendix B

CARS and DSM-IV Checklist Administration Instructions

- Step 1:** Introduce yourself to the informant
- Step 2:** Explain to staff that you will be asking a few questions about the client's behavior
- Step 3:** Find a quiet place in the home or office that is free of distractions
- Step 4:** Sit down next to the staff and open the CARS to the first page
- Step 5:** Place the open CARS on the table between you and the staff so that each of you is able to read the items
- Step 6:** Begin with Item 1.

"I want to know about how (client's name) relates to other people. I'm going to read you four statements. I want you to pick the statement that best describes (client's name) behavior. If you think that (client's name) is between two descriptions, you can give him/her a rating of 1.5, 2.5, or 3.5. Here are the statements about how (client's name) relates to people."

Read the number and entire description for each of the statements for Item 1, allowing the staff to follow along as you read. The scale must be read verbatim.

- Step 7:** Once you have read the statements that correspond to each item, ask the staff to choose the number/statement that best describes how (client's name) relates to other people. Circle the number on Item 1 and proceed to Item 2.
- Step 8:** Complete Items 2-14 in the same manner as Item 1. Read the descriptions for each number verbatim off the scale, allowing the staff to follow along as you read.
- Step 9:** For Item 15 read the descriptions for each number, allowing the staff to follow along as you read. Ask the staff to choose the number/statement that best describes the staff's general impressions about (client's name). If the staff indicates that (s) he does not know what autism is, state, "Just do your best to rate (client's name) on this scale". Circle the number and proceed to the DSM-IV checklist.
- Step 10:** Administer the DSM-IV checklist to the staff. Read the following statement:

"I have a few more questions to ask you about (client's name) behavior. I'm going to read you 12 descriptions of different behaviors and I want you to tell me yes or no whether each item describes (client's name)."

Read each of the 12 items on the DSM-IV checklist to the staff. After each item ask the staff whether this statement describes (client's name) behavior. If the staff says a behavior only happens to a certain degree (i.e. sometimes, every once in awhile, every now and then), the item should be scored as a "yes".

Vita

Noha Farrah Minshawi was born in Columbus, Ohio. Noha went to The Ohio State University and received a Bachelor of Science with Honors and Distinction in psychology in 2001. She is currently enrolled in the clinical psychology doctoral program at Louisiana State University, working under Dr. Johnny L. Matson. Her research and clinical area of interest is individuals with mental retardation and developmental disabilities.