A VALIDITY STUDY OF THE CHILDHOOD AUTISM RATING SCALE WITH AUTISTIC ADOLESCENTS: TWO METHODS

THESIS

Presented to the Graduate Council of the North Texas State University in Partial Fulfillment of the Requirements

For the Degree of

MASTER OF SCIENCE

By

Susan B. Rebillet, B.A.

Denton, Texas

August, 1986

Autistic symptoms appear to change during adolescence. This study replicates an earlier one (McCallon, 1985) which attempted to validate the use of the Childhood Autism Rating Scale (CARS) with adolescents. The standard observation method of administration was used. Additionally, information obtained in structured parent interviews was used to score a CARS. Subjects were nine autistic and nine nonautistic retarded adolescents from 13 to 21 years of age.

Results suggested that the CARS, administered either way, validity discriminates between retarded adolescents who are diagnosed as autistic and those with other handicaps. Methodological limitations of the present study are discussed.
TABLE OF CONTENTS

LIST OF TABLES ........................................ iv

A VALIDITY STUDY OF THE CHILDHOOD AUTISM RATING SCALE
WITH AUTISTIC ADOLESCENTS: TWO METHODS

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

The Childhood Autism Rating Scale (CARS)
Parent Interview Methods
Autism in Adolescence
  Behavior
  Language and Cognitive Development
  Social Adjustment
  Affective Development
Purpose of the Study

Method ................................................. 16

Subjects
Materials
Procedure

Results .................................................. 20

Discussion .............................................. 24

Appendix ............................................... 30

References ............................................. 36
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Subject Characteristics</td>
<td>17</td>
</tr>
<tr>
<td>2. Analysis of Variance for Group (Autistic versus Nonautistic) and Measure (Observed versus Parent Report)</td>
<td>21</td>
</tr>
<tr>
<td>3. Means and Standard Deviations of CARS Scores by Group and by Method</td>
<td>22</td>
</tr>
<tr>
<td>4. Item by Item Analysis for Groups and Methods</td>
<td>23</td>
</tr>
</tbody>
</table>
A VALIDITY STUDY OF THE CHILDHOOD AUTISM RATING SCALE WITH AUTISTIC ADOLESCENTS: TWO METHODS

The diagnosis or classification of children as autistic has been the subject of intense investigation for approximately forty years now, ever since the syndrome was identified by Leo Kanner (1943). Several diagnostic tools/scales/measures have emerged making the task of mental health workers and educators far more exact (Freeman, Guthrie, Ritvo, Schroth, Glass, & Frankel, 1979; Kurg, Arick, & Almond, 1980; Rimland, 1971; Ruttenberg, Dratman, Franknoi, & Wenar, 1966; Schopler, Reichler, DeVellis, & Daly, 1980). However, characteristics of the autism syndrome appear to change after puberty in ways which may make the use of instruments which are valid and reliable for children under twelve years of age less so for adolescents (Mesibov, Schopler, & Sloan, 1983). In fact, until recently we have had no empirical evidence as to whether the changes noted by many investigators were in fact characteristic enough of autistic adolescents to differentiate them from younger children with autism on a rating scale (McCallon, 1985).

Autism is presently defined as a syndrome characterized and diagnosed by a cluster of behaviors (Rutter, 1978). Diagnosis and differentiation of autism from other
disturbances is difficult because exact etiology is essentially unknown (Coleman, 1980; Folstein, 1980; Ritvo, 1980; Rutter, 1980). However, some definitions of autism are widely accepted in the current literature and serve as a basis for classification. From a review of the literature, there is general agreement on age at onset (before 30 months), impaired developmental rates and/or sequences, disturbances in language and cognitive abilities, and impaired capacity to relate to people, events, and objects (American Psychiatric Association, 1980; National Society for Autistic Citizens (NSAC), 1978; Ritvo & Freeman, 1978; Rutter, 1978). Other frequently cited criteria are insistence on sameness and disturbances in response to sensory stimuli (American Psychiatric Association, 1980; Rutter, 1978).

Various diagnostic instruments and screening devices exist which purport to identify autistic children and to discriminate them from other handicapping conditions. Some have adequate reliability and validity (McCallon, 1985; Parks, 1982; Teal, 1981).

One such instrument, the Childhood Autism Rating Scale (CARS; Schopler et al., 1980), will be the focus of the present investigation. The technical aspects of the CARS will be discussed. Parent interview methods will be reviewed with a view towards using the CARS with parent reports. Finally the importance of research with adolescent samples will be presented.
The Childhood Autism Rating Scale (CARS)

The Childhood Autism Rating Scale (CARS), see Appendix A, is an instrument that is reported by its authors to have adequate internal reliability (.94) as well as interrater reliability (.71) (Schopler et al., 1980). The authors obtained support for criterion validity of the CARS by comparing CARS scores to clinician's ratings of psychosis during the observations and to independent clinical assessments of autism by child psychologists and psychiatrists. The correlations were .84 and .80, respectively. These investigators report that the CARS discriminated between children with autism and children with other developmental disabilities but give no empirical data to support this.

Subsequent studies have reported excellent discriminant validity in differentiating children with autism from those with other handicapping conditions, principally mental retardation. Correlation coefficients are usually in the high .80s (McCallon, 1985; Teal, 1981). Teal compared the CARS to Rimland's Form E-2, a parent report questionnaire, and another instrument, the Autism Screening Instrument for Educational Planning (ASIEP) in their ability to distinguish between a group of autistic children and a matched group of trainable mentally retarded children. She states, "...the ability of the CARS to separate the groups was with 100% accuracy while the best ASIEP predictors were 95% accurate"
(p. 66). And further, "The CARS was found to be the most predictive instrument in this study" (p. 68). The Form E-2 "was the least predictive of the three instruments evaluated." There was overlap between the groups with Form E-2 and it was judged an inappropriate instrument for purposes of classification. McCallon (1985) also supported the CARS reliability (alpha coefficients = .75 for autistic and .82 for nonautistic) and discriminant validity, $t(19 = 13.73, p < .05$, with an adolescent sample of autistic and matched nonautistic retarded subjects.

The CARS is a 15 item rating system which incorporates Kanner's primary features of autism with characteristics identified by other scales. The definition of autism proposed by Rutter (1978) and Ritvo and Freeman (1978) comprise the theoretical basis for the CARS. These definitions recognize that there may be great variation in individual manifestations of the syndrome and thus the CARS reflects and is sensitive to these individual variations. It has been used to evaluate 537 children over a period of ten years (Schopler et al., 1980).

The CARS is administered by having trained raters observe the children's behavior. Ratings are made on each item along a 7 point continuum from 1 to 4 with 1 indicating no abnormality and 4 indicating severe abnormality. The child's chronological age must be taken into account when rating behaviors. The peculiarity, frequency and intensity
of the behavior also influence the ratings and allow for individual variability.

Scores on the CARS range from 15 to 60. The cut-off point was established at 30 for preliminary classification of autistic. Scores over 36, including five or more individual items scored 3 or greater, were classified as severely autistic. A score between 30 and 36 is classified as mild/moderately autistic.

**Parent Interview Measures**

The use of parent reports of their children's behavior for diagnostic or classification purposes has not often been experimentally addressed. Given that parental reports of age at onset of symptoms and other biographical data is nearly always heavily weighed in a clinical diagnosis of autism it would seem that investigation of their usefulness would be in order.

There are limitations to observations of behavior. Autistic behavior is erratic and variable (Ritvo & Freeman, 1978; Rutter, 1978) making any one observation limited. Children's behavior in an unfamiliar, such as a testing, situation may not be typical of their behavior. Many observations of children in the public schools for classification purposes are done in their classrooms. The raters are limited in the amount of time they can spend observing the child. It seems logical to assume that parents, who have observed their children's behavior across
time, would be rich and reliable sources of information regarding the behaviors used to assess the presence/absence of autistic characteristics. Parental reports of behaviors would thus become an alternative to direct observations if this proves to be true. In addition, a structured interview for gathering the data would standardize techniques rather than relying on subjective informal interviews for this critical information. Mesibov (1985, note 1), a leading authority on the syndrome of autism, feels that parent reports can be used in the assessment process.

A review of various diagnostic instruments reveals only one, Rimland's Diagnostic Checklist for Behavior Disturbed Children, Form E-2 (1971), which relies on parent reports of their children's behavior and symptoms. The Form E-2 is an extensive questionnaire to be completed by the parents containing 80 items. It was designed to differentiate the classic form of autism as defined by Kanner (1943) from other interpretations of the syndrome, as well as from other handicapping conditions. It is a revision of Rimland's original questionnaire, Form E-1, which consisted of 76 items addressing birth history symptomatology, speech patterns, and age of onset. Because of information obtained from the initial response to Form E-1, necessary changes were made which resulted in Form E-2. The most notable change is that the E-1 was to be used to gather information up to age 7 and the E-2 covers only the first five years (Rimland, 1974).
Parents responses indicated a change in symptoms after age five and a half, at which time the Kanner classical symptoms became more varied and individualistic.

The items are scored by assigning one + point to each item characteristic of classical autism and one - point for each question answered that is not characteristic of classical autism. The total score is the difference between the + and - scores. Scores can range from -42 to +45 with a score of +20 or higher considered indicative of early infantile autism and scores below +20 to be noted as "autistic-type." The total score is a combination of a behavior score ranging from -33 to +37 and a speech score ranging from -9 to +13. These scores are based on the analysis of 2,218 E-2 forms. A construct validity study was done to demonstrate differentiation between classical early infantile autism (EIA) and "autistic-type" syndromes (Rimland, 1971). Items which were most indicative of EIA showed large percentage differences from the remaining items and did discriminate between these two groups.

However, as Parks (1983) reports, there is no empirical evidence supporting the reliability of the Form E-2. Furthermore, a study of discriminant validity (DeMyer, Churchill, Pontius, & Gilkey, 1971) with the Form E-2 failed to discriminate between early schizophrenia, higher functioning autism, and lower functioning autism. In the standardization sample for the CARS, the E-2 was used to
evaluate 450 of the children (Schopler et al., 1980). Only eight of them were classified as autistic. Three of those eight were not identified as autistic by the CARS. Other researchers have compared responses on the E-2 from parents versus teachers (Prior & Bence, 1974). A total of 63 discrepancies were noted between the items for the subjects. Of the sample of nine children, six were rated as being "more autistic" by the teachers than by the parents with the reverse being true in only two cases. It is not clear why the parents rated their children as less autistic than the teachers. Prior and Bence suggested that they were comparing their child's behavior in the present with what it had been in the past.

Thus, Rimland's Form E-2 appears relatively unsound psychometrically and should only be used for research purposes until better validity as well as reliability can be demonstrated. It is not clear whether it is the Form E-2 or parent reports per se which are not useful. Therefore, this study undertakes to examine the feasibility of using parental reports of behavior gathered in a structured interview with a different, psychometrically sound, instrument--the CARS.

**Autism in Adolescents**

Within the generally agreed upon criteria for a diagnosis of autism, changes in the syndrome during the little studied period of adolescence have recently been
documented (Schopler & Mesibov, 1983). Freeman, Schroth, Ritvo, Guthrie & Wake (1980) point out that "the importance assigned to any one behavior or symptom can and does vary with the mental and chronological age of the child" (p. 346). A valid question is, are measures which are initially designed for children appropriate for adolescents due to maturational/developmental changes in the syndrome?

What is known about these changes comes from the very few longitudinal and cross-sectional studies which compared children and adolescents with autism (Ando & Yoshimura, 1979; Ando, Yoshimura, & Wakabayashi, 1980; Bartak & Rutter, 1976; DeMyer, Barton, DeMyer, Norton, Allen, & Steele, 1973; Lotter, 1974; McCallon, 1985; Rutter, 1970; Rutter, Greenfeld, & Lockyer, 1967; Rutter & Lockyer, 1967). Most follow-up and cross-sectional studies address what changes take place in adolescence and early adulthood by reporting different salient childhood variables and correlating them with outcome. The single most important prognostic factor in all the above cited studies was initial IQ. There appear to be some differences in symptom patterns, certain phenomena associated with autism, and outcome measures between autistic children with IQs over and under 50, 60, or 70 in these studies. It is estimated that 5 to 10 percent of autistic children have average range IQs, 12 to 20 percent are mildly retarded, and 43 to 68 percent are moderately to profoundly retarded (DeMyer, Norton, & Barton, 1973).
For an objective diagnosis it is essential to know which behaviors observed are unique to autism and which are the result of simple biological maturation or delayed development. It is well documented in the literature that the great majority of autistic children are also retarded (Brooker & Mareth, 1982; DeMyer, Barton, DeMyer, Norton, Allen, & Steele, 1973). Thus a matched group of mentally retarded controls would be essential in validating any classification measure. Ando & Yoshimura (1979), DeMyer et al. (1973), and DeMyer, Norton, and Barton (1973) all have used matched mentally retarded controls in their studies.

The findings from these various longitudinal and cross-sectional studies will be discussed as they pertain to the CARS. It should be noted that the changes to be discussed are basically relative improvements in a very severe disorder rather than improvements which bring normalcy. A recovery to normalcy is extremely rare (Rutter, 1978).

Behavior. Improvements consistently noted in lower functioning adolescents are primarily in self help behaviors of toileting, eating skills, self-control, and participation in group activities. However, these improvements do not seem to be a function of development or maturation but must be taught. Ando et al (1980) state that even though the adaptive behaviors of autistic persons improve with increasing age, children with autism are still less skilled
than mentally retarded controls. One study (DeMyer et al., 1973) showed little change in withdrawn or "loner" behavior except some 13 percent of the sample who developed immature, overfriendly behavior. Rutter and Bartak (1973) noted a general decrease in amount, although no change in character, of stereotyped behaviors, rituals, and self-injurious behavior. A marked change is from hyper- to hypo-activity if the child was hyper-kinetic at initial evaluation (Ando & Yoshimura, 1979; Rutter et al., 1967). One study (Gillberg & Schaumann, 1981) points to cases of severe deterioration in functioning along with the loss of skills previously attained when these children became adolescents.

Behavior which showed no changes with age in the lower functioning group were rocking, destruction of property, aggressive behavior, initiative, resistance to change, and attachment to odd objects (Ando & Yoshimura, 1979; Ando et al., 1980; DeMyer et al., 1973; Rutter et al., 1967; Rutter & Bartak, 1973). In the small percentage of higher IQ adolescents (IQ > 60) resistance to change has been noted to decrease (Rutter et al., 1967).

The implications of these findings for items on the CARS might be reflected in lowered ratings on (a) relationships with people and emotional responses due to improvements in self-control and participation in group activities, (b) use of body due to decreases in amount of
these stereotyped behaviors, and (c) fear or nervousness reactions given their increased self-control.

**Language and Cognitive Development.** The great majority of the lower functioning children in these studies had no speech at all upon initial evaluation but many had made progress at follow-up. DeMyer et al. (1973) report that 50 percent of their sample gained some communicative speech with some improvements in verbal IQ. However, these researchers also found an unexplained decrease in performance IQ with age. Bartak and Rutter (1976) note overall a much greater delay in speech in the low IQ group and in an earlier study (Rutter & Bartak, 1973) report that an IQ of less than 44 predicted that no phrase speech would be acquired with age. Ando and Yoshimura (1979) found significant relative improvements in comprehension and in conversation with age in their sample of children and adolescents (IQ < 50).

Echolalia and pronomial reversal seem to be two characteristics which do not change with age (Bartak & Rutter, 1976). DeMyer et al. (1973) noted that splinter skills of any sort which were present at the beginning of the study for the lower functioning children either disappeared or failed to continue to develop complexity at follow-up. Significant increases in cognitive skills (reading skills and arithmetic) were noted in the higher IQ (> 50) group (Rutter & Bartak, 1973).
The implications of these findings for items on the CARS might be reflected in lowered ratings on (a) talking due to improvements in conversation and communicative speech and, (b) inconsistencies in intelligence due to the disappearance or lack of development of splinter skills and the decrease in performance IQ.

Social Adjustment. Most investigators have noted a general developmental trend toward greater social responsiveness. When DeMyer et al.'s (1973) study began, all the children in the low IQ group were considered psychotically withdrawn, oblivious, or loners. At follow-up 75 to 80 percent remained in these categories. The areas most improved were eye contact, facial expression, and participation in play and group activities. Rutter, Greenfeld, and Lockyer (1967) state that autistic adolescents show little social "know-how," no awareness of others feelings, and commonly demonstrate socially embarrassing behavior. However, in the small portion of adolescents with IQs over 60, efforts to reach out to others have been noted (Bartak & Rutter, 1976; DeMyer et al., 1973).

The implications of these findings for items on the CARS might be reflected in lowered ratings on (a) visual attention and relationships with people due to increases in eye contact and participation in play and group activities, and (b) emotional responses due to improved facial expressions. On the other hand, ratings on emotional
responses may be higher due to their lack of awareness of socially appropriate behavior.

**Affective Development.** Very little has been written about the affective status of the autistic child probably because the symptoms often prevent the "usual" assessment of emotions. However, marked fear and nervous reactions have been noted in infancy and childhood and are among the features of the disorder showing the greatest variability with age (Ando et al., 1980; Rutter, 1967). Many low functioning children show no decrease at all in fears or tantrums at adolescence but many others (IQ > 50) show a marked decrease in fears and tantrums (Ando & Yoshimura, 1979; Rutter, 1968).

In summary, changes which may affect the scores on diagnostic instruments seem to be (a) the increases in self-help/adaptive behaviors, (b) the puzzling decrease in performance IQ in the lower functioning group, (c) the disappearance or lack of development of splinter skills, (d) the large decreases in fear and nervousness reactions in some adolescents, (e) the increase in social behavior such as eye contact and group participation, (f) the increase in self-control, (g) decreases in amounts of stereotyped, ritualistic, and self-injurious behavior, (h) improvements in conversation and communicative speech, and (i) improvements in facial expressions. Areas which appear to remain as disturbed as in childhood are (a) relationships with
others are almost never considered "normal," (b) speech abnormalities remain very problematic, (c) the adolescent remains very attuned to sensations--especially rocking sensations--although this does become less obsessive.

Thus, there is documented evidence of changes in the autistic syndrome in adolescence. It seems that these changes are individually variable and depend to some extent on IQ. From McCallon's (1985) study, it appears that the CARS, given its flexibility and sensitivity to variation, can incorporate these changes and still remain a reliable and valid instrument for preliminary classification of adolescents with autism.

Purpose of the Study

This study was undertaken to replicate the findings in support of the reliability and validity of the CARS for an adolescent population. It also attempted to validate clinical impressions and intuitive notions that parents are reliable reporters of their children's behavior. Autistic adolescents between 12 and 21 were matched with a group of nonautistic adolescents for IQ, age, and as closely as possible for sex and ethnicity. The CARS was administered both in the usual manner, through observation of the child, and through structured parent interviews. Hypotheses were as follows.

1. The CARS is not a reliable and valid instrument for use with adolescents.
2. Information elicited through a structured parent interview and subsequently used to obtain CARS scores are as useful as the observed CARS scores.

Method

Subjects

Subjects were nine retarded autistic adolescents and a matched group of nine nonautistic mentally retarded adolescents. All were students in a large southwestern metropolitan public school district and had been identified by the district's special education department. Refer to Table 1 for a description of the two samples.

The autistic adolescents comprised all adolescents 12 years and older identified by the school district as autistic or as atypical pervasive developmental disorder (7 and 2, respectively). The school district's diagnosis of autism was based on (a) a CARS rating of 30 or higher administered by the supervisor diagnosticians assigned to the schools, (b) autistic-like performance on the Adolescent and Adult Psychoeducational Profile, a diagnostic test for autistic people over 12 years old (Mesibov, Schopler, & Schaffer, 1985), (c) reported age at onset of less than 30 months, and (d) the clinical impression of the team which was led by a psychologist and included a speech therapist and various district special education personnel. The two students with atypical pervasive developmental disorder had the above characteristics except that reported age of onset was after 30 months.
Table 1

Subject Characteristics

<table>
<thead>
<tr>
<th></th>
<th>Autistic</th>
<th>Nonautistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>13-21</td>
<td>13-20</td>
</tr>
<tr>
<td>Mean</td>
<td>16.6</td>
<td>17.0</td>
</tr>
<tr>
<td>IQ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>10-67</td>
<td>12-62</td>
</tr>
<tr>
<td>Mean</td>
<td>35.0</td>
<td>34.3</td>
</tr>
<tr>
<td>SD</td>
<td>16.7</td>
<td>15.4</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>Female</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>White</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

The nonautistic group was matched for IQ, \( t(16) = .09, \ p > .05 \), and chronological age, \( t(16) = .34, \ p > .05 \), and as closely as possible for sex. Although the groups were not matched on ethnicity they were similar with 5 blacks and 4 whites in the autistic population and 3 blacks, 5 whites, and 1 hispanic in the nonautistic sample. The nonautistic
sample was selected from computer records of special education students in the district.

IQ scores were obtained from school records. Tests used to obtain these scores were the Wechsler Intelligence Scale for Children - Revised (3 scores), the Leiter International Scale (4 scores), the Merrill-Palmer Preschool Scales (2 scores), the Bayley Scales of Infant Development (2 scores), and the Stanford Binet Intelligence Scales (6 scores). Two of the scores for the autistic population were estimated by the school psychologist from the above tests and one "untestable" male was arbitrarily assigned a score of 10. The mean IQ of the autistic population was 35 with a range of 10 to 67 while the mean of the nonautistic sample was 34.3 with a range of 12 to 62. This sample of autistic students had IQ scores somewhat lower than those reported by DeMyer et al. (1973). She reported that between 43 and 68 percent of the autistic population was classified as moderately to profoundly retarded. In the present sample, 89 percent were so classified. Present IQ scores were also lower than the Janicki and Jacobson (1982) sample of 314 autistic adults which showed 85 percent with IQs in the retarded range.

Materials

The Childhood Autism Rating Scale (CARS; Schopler et al., 1980) was the principle instrument used to evaluate the groups (see Appendix A). It is a series of 15 scales which are scored from observation of the child's behavior
along a continuum from normal to severely abnormal behavior. The scoring criteria are empirically derived. Each of the items is scored from 1 (no abnormality) to 4 (severe abnormality) with half-point scores possible. A total score of 30 is the cut-off point for a preliminary diagnosis of autism.

A structured interview outline was used to elicit data for raters to use in scoring the parent interview CARS (see Appendix B). It consisted of a series of open-ended questions designed to obtain the greatest amount of information on which to base the CARS scores.

**Procedure**

Two psychology graduate students were trained to administer and score the CARS from observations of the subjects. Training consisted of (a) viewing training tapes (produced by the developers of the CARS) which allowed the graduate students to practice scoring a CARS, and (b) scoring a CARS along with trained school personnel in a different school district until each student could reliably administer the test ($r = .91$ and .90).

A letter was sent to all the parents of the autistic children and the possible "match parents" asking if they would be willing to participate in this research (see Appendix C). From the children's parents who returned the signed agreement to participate the best matches were made.

The CARS was administered to autistic and nonautistic subjects in their schools by the two raters. Subsequently,
the two raters went to the parents homes at their convenience and interviewed them with the structured interview outline, taped the interview, and then scored a CARS based on the information elicited. The raters interviewed mothers in 16 cases, a father who had custody in one case, and a grandmother and mother in one case. The raters were counterbalanced across children and parents, i.e., students interviewed parents of children they had not observed. During the observations and scoring of the CARS interrater reliability was assessed between the two students and found to be acceptable (r = .97). This reliability was based on observations of the same children. The audio tapes were used to obtain interrater reliability for the parent interview CARS (r = .95). This rater reliability was based on hearing and scoring the audio tapes of the parent interview done by the other student.

Results

This study was designed to (a) provide additional discriminant validity for the use of the CARS with the retarded autistic adolescent following McCallon (1985), and (b) to provide interrater reliability and validity information for the use of parent interview information to score a CARS. Findings related to discriminant validity will be reported first, followed by those related to parent interview CARS. An alpha level of .05 was adopted for all analyses.
A t test comparison between the school administered CARS scores and the CARS scores administered for this study showed a significant difference, $t(8) = 2.38, p < .05$. The school raters rated the children as a whole more severely autistic than the study raters but both groups of raters identified all the children as autistic, the only difference being the severity of the syndrome.

A 2 X 2 analysis of variance with group (autistic versus nonautistic) as the between subject factor and measure (observed versus parent interview) as the within subject factor was conducted. See Table 2 for ANOVA information. There was a significant main effect of group, $F(1, 16) = 1169.40, MSe = 2770.66$.

Table 2

<table>
<thead>
<tr>
<th></th>
<th>Sum Squares*</th>
<th>Mean Square</th>
<th>F</th>
<th>Tail Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>135669.4</td>
<td>135669.5</td>
<td>1169.40</td>
<td>.0000</td>
</tr>
<tr>
<td>Error</td>
<td>44330.5</td>
<td>2770.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Measure</td>
<td>711.1</td>
<td>711.1</td>
<td>.45</td>
<td>.5103</td>
</tr>
<tr>
<td>Measure X Group</td>
<td>2.8</td>
<td>2.8</td>
<td>.00</td>
<td>.9669</td>
</tr>
<tr>
<td>Error</td>
<td>25086.1</td>
<td>1567.9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*df = 1, 16.
CARS score means were 36.13 and 23.85 for the autistic and nonautistic groups, respectively. Table 3 reports means and standard deviations for the CARS scores by group and by measure.

Table 3
Means and Standard Deviations of CARS Scores by Group and by Method

<table>
<thead>
<tr>
<th></th>
<th>Autistic</th>
<th>Nonautistic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Observations</td>
<td>35.67</td>
<td>5.8</td>
</tr>
<tr>
<td>Parent-Interview</td>
<td>36.61</td>
<td>4.7</td>
</tr>
</tbody>
</table>

There was no main effect of measure; observed scores and parent interview scores were not significantly different from one another, $F(1, 16) = .45$, $MSe = 1567.9$. In addition, no interaction was found between groups and observed or parent interview CARS, $F(1, 16) = .00$, $MSe = 1567.9$.

Interitem consistency was assessed by calculating an alpha coefficient. For the autistic group, the alpha coefficient for the observed CARS was .83 and for the parent interview CARS .71. For the nonautistic group the alpha coefficients were .78 for observed CARS and .77 for parent interview CARS. An item by item analysis (see Table 4) indicated that within the autistic group, the item # 14, inconsistencies in intelligence was negatively correlated
Table 4

Item by Item Analysis for Groups and Methods

<table>
<thead>
<tr>
<th>Item #</th>
<th>Item Descriptor</th>
<th>Observed Autistic Item-Total Correlation</th>
<th>Alpha if Item Deleted</th>
<th>Parent-Report Item-Total Correlation</th>
<th>Alpha if Item Deleted</th>
<th>Observed Nonautistic Item-Total Correlation</th>
<th>Alpha if Item Deleted</th>
<th>Parent-Report Item-Total Correlation</th>
<th>Alpha if Item Deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Relate</td>
<td>-.19</td>
<td>.85</td>
<td>.50</td>
<td>.68</td>
<td>.89</td>
<td>.71</td>
<td>.22</td>
<td>.78</td>
</tr>
<tr>
<td>2</td>
<td>Imitate</td>
<td>.47</td>
<td>.82</td>
<td>.20</td>
<td>.71</td>
<td>.60</td>
<td>.75</td>
<td>.25</td>
<td>.78</td>
</tr>
<tr>
<td>3</td>
<td>Emotion</td>
<td>.57</td>
<td>.82</td>
<td>.35</td>
<td>.70</td>
<td>.40</td>
<td>.78</td>
<td>.72</td>
<td>.73</td>
</tr>
<tr>
<td>4</td>
<td>Use Body</td>
<td>.59</td>
<td>.83</td>
<td>.45</td>
<td>.68</td>
<td>.67</td>
<td>.75</td>
<td>.70</td>
<td>.72</td>
</tr>
<tr>
<td>5</td>
<td>Objects</td>
<td>.66</td>
<td>.80</td>
<td>.03</td>
<td>.73</td>
<td>.67</td>
<td>.74</td>
<td>.91</td>
<td>.71</td>
</tr>
<tr>
<td>6</td>
<td>Change</td>
<td>.56</td>
<td>.82</td>
<td>.09</td>
<td>.73</td>
<td>.67</td>
<td>.74</td>
<td>.91</td>
<td>.71</td>
</tr>
<tr>
<td>7</td>
<td>Visual</td>
<td>.53</td>
<td>.82</td>
<td>.68</td>
<td>.66</td>
<td>.79</td>
<td>.73</td>
<td>.15</td>
<td>.77</td>
</tr>
<tr>
<td>8</td>
<td>Hear</td>
<td>.66</td>
<td>.81</td>
<td>.72</td>
<td>.64</td>
<td>.33</td>
<td>.78</td>
<td>.65</td>
<td>.73</td>
</tr>
<tr>
<td>9</td>
<td>Touch</td>
<td>.66</td>
<td>.81</td>
<td>.75</td>
<td>.65</td>
<td>-.10</td>
<td>.82</td>
<td>.18</td>
<td>.77</td>
</tr>
<tr>
<td>10</td>
<td>Fear</td>
<td>.73</td>
<td>.80</td>
<td>.27</td>
<td>.71</td>
<td>.49</td>
<td>.77</td>
<td>-.06</td>
<td>.80</td>
</tr>
<tr>
<td>11</td>
<td>Talk</td>
<td>.46</td>
<td>.82</td>
<td>.01</td>
<td>.72</td>
<td>.79</td>
<td>.73</td>
<td>.69</td>
<td>.72</td>
</tr>
<tr>
<td>12</td>
<td>Point</td>
<td>.62</td>
<td>.81</td>
<td>.42</td>
<td>.68</td>
<td>.49</td>
<td>.77</td>
<td>.03</td>
<td>.78</td>
</tr>
<tr>
<td>13</td>
<td>Activity</td>
<td>.08</td>
<td>.85</td>
<td>.31</td>
<td>.70</td>
<td>-.11</td>
<td>.82</td>
<td>.18</td>
<td>.78</td>
</tr>
<tr>
<td>14</td>
<td>Intelligence</td>
<td>-.42</td>
<td>.85</td>
<td>-.15</td>
<td>.74</td>
<td>.49</td>
<td>.77</td>
<td>.34</td>
<td>.76</td>
</tr>
<tr>
<td>15</td>
<td>General</td>
<td>.69</td>
<td>.81</td>
<td>.58</td>
<td>.69</td>
<td>-.44</td>
<td>.83</td>
<td>.77</td>
<td>.74</td>
</tr>
</tbody>
</table>

\(^a\)Alpha level for autistic observed was .83.

\(^b\)Alpha level for autistic parent report was .71.

\(^c\)Alpha level for nonautistic observed was .78.

\(^d\)Alpha level for nonautistic parent report was .76.

\(^e\)Items which had zero variance for this group.
with total score. This was the case for both observed and parent interview CARS. For the observed CARS only, the item #1, relationships with people, was also negatively correlated with the total score. Within the nonautistic group, the item #15, general impressions, was negatively correlated with total score only for the observed CARS.

If the item inconsistencies in intelligence was deleted from the scale for the autistic group it would have raised the alpha level for the observed and parent interview CARS to .85 and .74 respectively from .83 and .71. Elimination of the relationships to people item for the observed CARS would raise the alpha level to .85 from .83. If the item, general impressions was deleted from the nonautistic group observed CARS that alpha level would increase to .82 from .78. Refer to Table 4 for complete details on the item by item analysis.

Discussion

The results of this replication study indicate support for McCallon's (1985) findings that the CARS is an appropriate instrument, both reliable and discriminantly valid, for use with retarded adolescents ages 12 to 21. The technical adequacy of the CARS seems to be relatively insensitive to age related changes in autistic characteristics.

In addition, the results indicate that parent interviews can be used to obtain a CARS rating of autism. It further shows that the parent interviews are valid when used to
discriminate their children with autism from other adolescents who are nonautistic.

Perhaps Rimland's Form E-2, which uses direct parent reports, is not as psychometrically sound as the CARS as DeMyer et al. (1971), Parks (1983), Prior and Bence (1974), and Schopler et al. (1980) have concluded. However, this does not therefore invalidate parent reports per se. In this study, structured parent interviews were used to elicit the information used to score a CARS. This proved to be as helpful as observation by trained raters. Furthermore, these results tend to support clinical impressions that parental reports correlate well with diagnosis of autism in adolescence. In fact, the phi correlation was 1.0 for parent interview CARS scores, i.e., none of the adolescents diagnosed autistic were given a score less than 30 on the parent interview CARS and none of the adolescents diagnosed nonautistic were given a score over 30.

Interitem reliability of the CARS for both groups and both types of CARS scores appears adequate (refer to Table 4). However, elimination of two items for the autistic group and one item for the nonautistic group would increase the alpha levels however slightly. The item, inconsistencies in intelligence, which was negatively correlated with total score for both parent interview and observed CARS within the autistic group may mean that this item needs revision.
This was one of Kanner's primary characteristics of childhood autism and since using his criteria alone has been demonstrated to identify fewer children as autistic than many other instruments, this item may have limited application to the wider definitions of autism presently in use (Cohen et al., 1978; Kanner, 1962, cited in Rimland, 1971; Parks, 1983). Another consideration may be the DeMyer et al. (1973) observation of the disappearance or lack of development of splinter skills which were seen in their sample at initial evaluation. The inconsistencies in intelligence may actually be prevalent in childhood but not in adolescence. One other hypothesis is that this item is a difficult item to score for those untrained in evaluating children's progress in school subjects and educational levels. Scorers do not usually have enough information to score this item accurately.

Elimination of the item relationships with people on the observed CARS was shown to raise the alpha level. This may indicate that adolescents, indeed, as the literature would suggest, do appear to make gains in social and adaptive behaviors across IQ differences. Trained raters are accustomed to observing severe nonrelating behavior on the part of autistic children and therefore may have been generous when there was a lessening of the severity. Perhaps even more interesting is that this negative correlation was not the case for the parent interview CARS. The information
gathered from them caused the raters to give scores which correlated with the child's overall CARS scores indicating that parents may perceive their children more accurately since they have the advantage of seeing their children in all circumstances.

With regard to the item negatively correlated with total score for the nonautistic group (general impressions), it may reflect the reluctance of raters to give a score of 1 (no abnormality) and 2 (mildly abnormal). A score of 3 states that the rater feels the child is moderately abnormal and thus, with many retarded adolescents who often have some bizarre characteristics, the rater is inclined to rate the child two and a half to three which is sometimes higher than they have scored on any other scale. The criteria for the general impressions category are the least explicit and therefore most subject to individual interpretations.

The CARS appears robust enough to weather the few negative correlations without resorting to revision of the measure which would require restandardization. McCallon (1985), with a larger sample, obtained negative correlations only for the item, inconsistencies in intelligence. If anything, this item might be eliminated or, at least, criteria for rating should be made more explicit. On the whole, however, the CARS performs its classification and/or screening duties admirably.

A limitation of this study is its small sample size. With a larger sample, the alpha coefficients would very
likely have been more acceptably and consistently in the 80s than in the 70s. A confound of this study is that the CARS, although administered by persons not connected in any way with this study, was used in conjunction with other information in the initial selection of subjects for this study.

The fact that all the adolescents in both groups had IQs under 68 and the small sample size precludes any meaningful analysis of CARS scores varying with IQ as McCallon (1985) had examined. Her speculation that sample selection may be biased; that is, that mildly autistic adolescents with mild retardation have adjusted to school routines and were therefore never referred to the newly created autistic program in the school district may very well be operating in this case. This district's organized autism program is quite new as well. On other study (Janicki & Jacobson, 1982) also suggests that this selection bias may exist in their sample of autistic adults. Thus, the question of the identification of mildly retarded or borderline retarded and mildly or moderately autistic adolescents using the CARS has yet to be addressed. Furthermore, as mentioned in the introduction, children with IQs in the higher ranges may evidence a different pattern than the majority of autistic children and thus these studies taken together would not seem to provide enough information for a judgement of reliability and validity for adolescents who are mildly
retarded (IQ scores from 52-67) or borderline (IQ scores from 68-83) since their numbers were so small.

In summary, this study supports the validity and reliability of the CARS as an assessment and classification instrument through direct observation. It gives support for the standardization of parent interviews as an alternative to direct observation when used to score a CARS. The CARS reliably identifies those who have been diagnosed autistic and it discriminates them from adolescents with other handicapping conditions. Structured parent interviews, if used to score a CARS, were shown to be as valid as the observed CARS, with accompanying adequate interrater reliability. An important unanswered question is how the observed scores and parent interview scores would behave with a sample of mildly retarded or borderline mildly retarded autistic adolescents. The samples have not been large and varied enough to bring evidence to bear on this question. It seems that empirical evidence for differential changes in the syndrome of autism for different IQ groups in adolescence appears to warrant further investigation, especially for the over 50 to 70 IQ group, since this study and others address primarily the under 50 to 70 group.
Appendix A

Childhood Autism Rating Scale (CARS)

I. RELATIONSHIPS WITH PEOPLE

(1) No evidence of difficulty or abnormality in relating to people.
(2) Mildly abnormal relationships.
(3) Moderately abnormal relationships.
(4) Severely abnormal relationships.

II. IMITATION (verbal and physical)

(1) Age-appropriate imitation.
(2) Mildly abnormal imitation.
(3) Moderately abnormal imitation.
(4) Severely abnormal imitation.

III. EMOTIONAL RESPONSES

(1) Age-appropriate and situation-appropriate emotional responses.
(2) Mildly abnormal emotional responses.
(3) Moderately abnormal emotional responses.
(4) Severely abnormal emotional responses.

IV. USE OF BODY

(1) Age appropriate use of body.
(2) Mildly abnormal use of body.
(3) Moderately abnormal use of body
(4) Severely abnormal use of body
Appendix A--Continued

V. RELATION TO OBJECTS

(1) Age appropriate use and interest with toys and in other objects.
(2) Mildly inappropriate interest or use.
(3) Moderately inappropriate interest or use of objects.
(4) Severely inappropriate interest or use.

VI. ADAPTATION TO ENVIRONMENTAL CHANGE

(1) Age appropriate resistance to change.
(2) Mild resistance to change.
(3) Moderate resistance to change.
(4) Severe resistance to change.

VII. VISUAL ATTENTION

(1) Age appropriate visual responses.
(2) Mildly abnormal visual responses.
(3) Moderately abnormal visual responses.
(4) Severely abnormal visual responses.

VIII. LISTENING ATTENTION

(1) Age appropriate listening responses.
(2) Mildly abnormal listening responses.
(3) Moderately abnormal listening responses.
(4) Severely abnormal listening.

IX. TOUCH, TASTE, AND SMELL

(1) Normal response to pain, smell, taste and touch.
(2) Mildly abnormal responses to pain, smell, taste or touch.
(3) Moderately abnormal response to pain, smell, taste or touch.
(4) Severely abnormal response to pain, smell, taste or touch.

X. FEAR OR NERVOUSNESS
(1) No excessive fear or nervousness.
(2) Mild fear or nervousness.
(3) Moderate fear or nervousness.
(4) Severe fear or nervousness.

XI. TALKING
(1) Speech is appropriate for age.
(2) Speech is mildly abnormal.
(3) Speech is moderately abnormal.
(4) Speech is severely abnormal.

XII. POINTING AND GESTURING
(1) Use of gestures is appropriate for age.
(2) Use of gestures is mildly abnormal.
(3) Use of gestures is moderately abnormal.
(4) Use of gestures is severely abnormal.

XIII. ACTIVITY LEVEL
(1) Activity level is normal.
(2) Activity level is mildly abnormal.
(3) Activity level is moderately abnormal.
(4) Activity level is severely abnormal.

XIV. INCONSISTENCIES IN INTELLIGENCE
(1) Intelligence is normal.
(2) Intelligence is mildly abnormal
(3) Intelligence is moderately inconsistent.
(4) Intelligence is severely inconsistent.
XV. GENERAL IMPRESSIONS
Appendix B

Structured Parent Interview Questions

1. Discipline-Behavior Problems
   Does your child listen to you?
   What behavior problems have you had?
   How has this changed?

2. Emotional Reactions
   How does he/she react to changes? Vacations? Moves?
   Changes in plans?
   Are weekends harder?
   How does he/she react to new places? Going to the store?

3. Relating to people
   Tell me what your child is like around most people.
   Is he/she different with one person versus a group?
   Does he/she have friends?
   How does he/she react to strangers?

4. Play
   What does he/she do during his/her free time?
   What things does he/she play with?
   What are his/her favorite things?
   If you gave your child a new object what would he/she do with it?

5. Imitation-Language
   How do you know that your child understands what you say?
   How does your child let you know what he/she wants or how he/she is feeling?
   Is there anything different about the way he/she talks?

6. Hearing or Looking
   What kinds of sounds does your child really like? Anything unusual?
   Do some sounds upset or disturb him/her?
   Does he/she look at things or people in unusual ways?
   How so?

7. Intelligence
   Does your child seem really smart in some areas but has problems in others?

8. Imitation
   Does your child imitate others? Or you?

9. What do you like best about your child? What things worry you the most?

10. What has changed the most for the worse? And for the better?
Dear Parent:

The Special Education Department of the Fort Worth I.S.D. and North Texas State University are now conducting a study in the Fort Worth Public Schools which will eventually help us to plan services for your child. A part of the study includes a parent interview. This part is very important and we need your cooperation very much. The interview would last only one-half to one hour and would be completely at your convenience. It will consist of talking about what your child's behavior is like.

We sincerely hope that you will decide to participate and the study will be more fully explained during the interview. Please return this letter in the stamped envelope enclosed to let us know if you will be willing to help us. Thank you very much.

Sincerely,

Susan B. Rebillet, Researcher
North Texas State University

Enclosure

____ Yes, I am willing to participate. Parent(s) of
____ No, I am not willing.

Approved: Assistant Superintendent for Secondary Education.
References


