A FOLLOW-UP STUDY OF AUTISTIC
AND AUTISTIC-LIKE CHILDREN

DISSERTATION

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

For the Degree of

DOCTOR OF PHILOSOPHY

By

Denise McCallon, M.S.
Denton, Texas
August, 1988

Autism is a lifelong handicapping disorder that occurs on a continuum of severity. Children who show mild autistic behaviors but do not meet the criteria for a diagnosis of autism are often called autistic-like, but it is not known if their development and functioning are similar to that of autistic children. A follow-up study was done on 35 autistic and autistic-like children who were an average of 3 years of age when initially seen. Initial test scores indicated that the children were similar on measures of intellectual/developmental functioning, receptive vocabulary, and adaptive functioning. Approximately 4 years later they were evaluated again. Using the Childhood Autism Rating Scale, the children were divided at follow-up into three groups: nonautistic, mildly/moderately autistic, and severely autistic. Most children made gains on intelligence tests and displayed a diminishing number of autistic symptoms. Changes in nonverbal intelligence, adaptive functioning and receptive vocabulary scores depended on group membership. The results are discussed in relation to the reported stability of cognitive functioning in young autistic children and the implications for clinical
practice, early intervention, and research on attachment. The nature of the syndrome of autism is also discussed, particularly in its relation to the milder, atypical children. The superior follow-up status of the autistic-like as compared to the autistic children raises serious questions about including the two groups in the same syndrome.
A FOLLOW-UP STUDY OF AUTISTIC AND AUTISTIC-LIKE CHILDREN

CHAPTER

I. Introduction ............................. 1

Problems With Definition
Differentiating Autism From Schizophrenia
The Issue of Mental Retardation
Summary
Problems With Objectivity of Measurement
Measurement of Severity
Measurement of Intelligence
Measurement of Outcome
Summary
Subclinical Autism: Children Who Are Autistic-Like
Summary
Purpose of the Study

II. Method ................................. 30

Subjects
Materials
Childhood Autism Rating Scale (CARS)
Autism Checklist
Problem Checklist and Parent Questionnaire
Vineland Social Maturity Scale
Tests of Intelligence/Developmental Functioning
Language Assessment
Outcome Rating
Procedure
Collection of Initial Data
Collection of Follow-Up Data
Reliability of Diagnosis
Combining Samples

III. Results ............................... 49

Comparisons Over Time
Comparisons Between Groups
Correlation Among Measures
TABLE OF CONTENTS—Continued

Prediction of Outcome

Summary

IV. Discussion ........................................ 66

Time-Related Changes

Autistic Behaviors
Cognitive Functioning
Adaptive Functioning
Behavior Management

The Severity Continuum
Findings of Infant Characteristics
Summary

Appendix .................................................. 93

References .............................................. 119
<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Subject Characteristics</td>
<td>31</td>
</tr>
<tr>
<td>2. Initial Tests of Intelligence/Developmental Functioning</td>
<td>40</td>
</tr>
<tr>
<td>3. Discrepancies in Diagnosis</td>
<td>44</td>
</tr>
<tr>
<td>4. Diagnostic Categories—Current</td>
<td>46</td>
</tr>
<tr>
<td>5. Comparison of Two Samples of Autistic Children</td>
<td>47</td>
</tr>
<tr>
<td>6. Initial and Follow-Up Evaluation Scores</td>
<td>50</td>
</tr>
<tr>
<td>7. Distribution of IQ Scores</td>
<td>51</td>
</tr>
<tr>
<td>8. Frequency of Displayed Symptoms From the Autism Checklist</td>
<td>52</td>
</tr>
<tr>
<td>9. Most Frequently Checked Problems From the Problem Checklist</td>
<td>53</td>
</tr>
<tr>
<td>10. Percentage Indicating Problems on Parent Questionnaire</td>
<td>54</td>
</tr>
<tr>
<td>11. Initial Clinic Diagnoses by Group</td>
<td>56</td>
</tr>
<tr>
<td>12. Mean Scores on Initial Measures by Group</td>
<td>57</td>
</tr>
<tr>
<td>13. Differences Among Groups on Follow-Up Measures</td>
<td>58</td>
</tr>
<tr>
<td>14. Distribution of Follow-Up IQ Scores by Group</td>
<td>59</td>
</tr>
<tr>
<td>15. Changes in SQs by Group</td>
<td>60</td>
</tr>
<tr>
<td>16. Change in Number of Autistic Symptoms</td>
<td>60</td>
</tr>
<tr>
<td>17. Work/School Outcome Rating by Group</td>
<td>62</td>
</tr>
<tr>
<td>18. Correlations Among Follow-Up Measures</td>
<td>63</td>
</tr>
<tr>
<td>19. Correlations Between Initial and Follow-Up Measures</td>
<td>64</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

Autism is a lifelong handicapping disorder which has a prevalence rate of less than 1 per 1,000 (Achenbach, 1982). The disorder is identified by its symptoms; however, there is no universal consensus regarding which symptoms are essential for a diagnosis. Authorities generally agree that autism is a rare pervasive developmental disorder in which three features must be present: a) impaired social development in relating to people, objects, and events, b) disturbance of language and cognitive skills, and c) early onset before 30 months of age (American Psychiatric Association, 1980; Ritvo & Freeman, 1978; Rutter, 1978). In addition, mental retardation is frequently associated with autism. Approximately 40 percent of autistic children have IQs below 50, and 70 percent have IQs below 70 (American Psychiatric Association, 1980). Currently the etiology is believed to be a physical dysfunction within the central nervous system, the exact nature and location of which has yet to be identified (Ritvo & Freeman, 1978). Both biological and genetic factors have been implicated (Schopler, 1983).

Kanner was the first to describe autism as a unique disorder. In a now classic 1943 article, he reported 11
case studies of children who had shown an extreme disinterest in people from the very beginning of life, an unusual attachment to objects rather than people, and a "powerful desire for aloneness and sameness" (p. 248). All the children had come from highly intelligent parents. He called the syndrome early infantile autism, and he recognized the importance of follow-up with this population of children.

Fortunately, many researchers were interested in following the development of children like those identified by Kanner. Follow-up studies have provided much of what we know about autism today (e.g., DeMyer et al., 1973; Eisenberg, 1956; Lotter, 1974b; Rutter, Greenfeld, & Lockyer, 1967). These studies have revealed that certain symptoms of the disorder change with age. For example, resistance to changes in the environment may become less pronounced as the autistic child grows (Rutter, Greenfeld, & Lockyer, 1967). Even the hallmark symptom of a lack of interest in relating to people can diminish; and an autistic person may reach out for social contact during the adolescent years (Rutter, 1970). They remain, however, a severely handicapped group throughout life. While a wide range of outcomes have been reported by follow-up studies, the prognosis for the vast majority of autistic children is quite poor (Lotter, 1978).

When these landmark follow-up studies were undertaken,
the definition of autism was not quite the same as it is today. The definition changed as more was learned about the disorder through empirical study. In addition, in defining any disorder, some children are of necessity excluded. There is a sizeable group of children who share characteristics of autism, specifically deficits in language and social development, not all of whom meet the recognized criteria for diagnosis. These children have been called autistic-like, but there is no consensus regarding their diagnosis. Researchers, concerned with accurately specifying the population they are studying, exclude the difficult-to-diagnose children. The result has been that, while autistic-like children exist in numbers perhaps greater than autism itself (Wing, 1978), it is not known where they should be classified or how their prognosis is different from that of autistic children.

The purpose of this study is to provide a follow-up evaluation of autistic and autistic-like children who were seen during the last 15 years at a children’s outpatient clinic. Due to the developmental changes which can occur, and the lifelong nature of the handicap of autism, follow-up is of critical importance. With respect to autistic children, previous follow-up studies have been hindered by significant problems (Balow, 1981; Freeman & Ritvo, 1981). Regarding autistic-like children, follow-up studies are lacking.
Among the problems associated with follow-up studies have been the lack of clearly defined initial populations due to varying definitions and criteria for the diagnosis of autism, and the lack of objective measurement of both initial and follow-up variables. Such problems affect our ability to confidently generalize their findings to autistic children today. First, the problems which have plagued follow-up studies of autistic children will be reviewed in order to specify the ways in which they can be minimized in the present study. Last, the literature that is available on autistic-like children will be reviewed.

Problems With Definition

Confusion about the definition of autism has been the rule rather than the exception. Authorities do not agree on which symptoms are essential for a diagnosis of autism. The search for the essential symptoms began almost immediately with Kanner. Subsequent to identifying the disorder, he attempted to define the essential symptoms. He believed that the extreme self-isolation and the obsessive insistence on sameness were primary characteristics, and the other associated behaviors were secondary (Kanner & Eisenberg, 1955). According to Rutter (1978), Kanner's reduction of the essential symptoms to these two resulted in confusion. The simplified criteria omitted the peculiar language deficits which had characterized all of Kanner's cases. Other researchers followed the lead in attempting to specify
the essential features, but unfortunately there was little agreement among them.

In retrospect, other major sources of confusion can be identified. Empirical research was challenging original assumptions. One of these assumptions was that autism was related to schizophrenia. Autism was erroneously linked with schizophrenia from the beginning by Kanner's choice of words. Prior to his use of the term, autism had referred to an "active withdrawal into fantasy shown by schizophrenic patients" (Rutter, 1978, p. 2). Therefore, it remained for autism and childhood schizophrenia to be differentiated. Another of these assumptions was that autistic children had average intellectual potential. It gradually became evident that autism could occur at all levels of cognitive functioning, but that it occurred most often at the retarded levels.

Differentiating Autism From Schizophrenia

Originally Kanner believed that autism was the earliest form of schizophrenia, and as a result many clinicians and researchers did not attempt to differentiate the two disorders (Schopler, 1983). Thus the early literature is a mixed "bag" of studies involving children described almost interchangeably as psychotic, autistic, and schizophrenic (for reviews see Achenbach, 1982; Lotter, 1978). Kanner did note ways in which the two disorders were different; namely, autism is characterized by an extreme aloneness from the
very beginning of life, as opposed to a change from relatively normal functioning which generally precedes schizophrenia. He also saw the purposeful relation to objects and the extreme desire for sameness as different from schizophrenia (Kanner, 1943). Unfortunately, the term childhood schizophrenia became a synonym for childhood psychosis, and for many workers in the field, it became a synonym for infantile autism as well (Achenbach, 1982).

In an attempt to clarify the confusion among investigators, a British Working Party proposed a nine-point criteria for diagnosis of what they called "early childhood psychosis or schizophrenic syndrome of childhood" (Creak, 1963, p. 84). What they were describing was autism (Lotter, 1966; Rutter, 1978) (see Appendix A). During the 1960's investigators adopted these criteria and the term childhood psychosis (Freeman & Ritvo, 1981). The Creak criteria, however, had its shortcomings. It failed to specify how many of the nine points were necessary to establish diagnosis (Rutter & Lockyer, 1967), and several of the points were "loosely described and were open to widely differing interpretations" (Schopler & Rutter, 1978, p. 508).

The criteria also failed to specify age of onset, resulting in the grouping together of psychoses beginning in early infancy with those not emerging until later childhood, or even adolescence. It was not until the late 1960's and
early 70's that research began to shed light on essential differences between these early and late onset psychoses (Rutter, 1978). A bipolar distribution of age of onset for psychoses was found, with one peak in infancy and one in adolescence. Children with early and late onset psychoses differed in terms of social class, family history of schizophrenia, evidence of cerebral dysfunction, and level of intelligence. Studies also found that autistic individuals do not develop delusions and hallucinations, and autism is not characterized by periods of remission, as is schizophrenia (Green et al., 1984; Rutter, Greenfeld, & Lockyer, 1967). In addition, the prognosis for autistic children appears, on the average, much worse than for children with schizophrenia. Whereas autism is a handicap which lasts a lifetime, approximately 20% of children with childhood schizophrenia may recover completely (Eggers, 1978).

Thus the process of differentiating autism from childhood schizophrenia led to findings which clearly have implications for current researchers. Autistic subjects should be precisely identified by recognized criteria which include early onset and the absence of hallucinations. Both of these are included in the criteria for the diagnosis of infantile autism adopted by the American Psychiatric Association for the third edition of its Diagnostic and Statistical Manual of Mental Disorders (DSM-III) (see
Appendix B).

The Issue of Mental Retardation

Kanner's original cases were an unusual group of children. Despite severe deficits in adaptive functioning, physically they did not appear to be retarded and some demonstrated extraordinary feats of rote memory (e.g., reciting the Presbyterian catechism) or adeptness with formboards and puzzles far beyond their other cognitive skills. For Kanner, and others who closely followed his lead, these unusual peak skills suggested normal intellectual potential. Subsequently, at least average potential was considered a criterion, even to the point of excluding retarded children from a diagnosis of autism (e.g., Rimland, 1971). This led to problems with the definition of autism because empirical research began to cast doubt on the assumption of average intellectual potential.

In the early literature many autistic children were called "untestable" because of the difficulties examiners had in maintaining standard testing procedures. Autistic children failed to respond to test materials, had very short attention spans, displayed extreme noncompliance, and even resorted to severe emotional outbursts. It is probable that these negative behaviors occurred because projective and other language-dependent tests were being used (Schopler, 1983). The children were being given material that was much too difficult for them. Rather than dispelling the myth of
average intellectual potential, however, the "untestability" of these children allowed the myth to continue unchallenged for some time.

Eventually it became apparent that, for all practical purposes, many children were retarded as well as autistic. Lockyer and Rutter (1969) found untestability to be simply a function of very low levels of intelligence. When Alpern (1967) gave test items at low levels of difficulty (infant levels) to "untestable" children, the subjects became testable and important data was obtained. With more realistic testing of these children (i.e., not language-based and not at levels too difficult for them to understand), the distribution of IQ scores among autistic children was found to be approximately 60 percent in the moderate-to-profound range of retardation, 20 percent in the mild range, and 20 percent in the borderline-to-normal range (Ritvo & Freeman, 1981; Schopler, 1983).

While most autistic children are retarded to some degree, a small subgroup is not. Apparently this was the group initially identified by Kanner. The presence of unusual peak skills appears to exist in a readily identifiable subgroup of higher functioning autistic children. In a large sample of autistic children, Schopler (1983) found that most who had peak skills were in the mildly retarded to normally intelligent range. Others have found subgroups of children, matching Kanner's syndrome, and
sharing many characteristics with a wider population of autistic children (Prior, Perry, & Gajzago, 1975; Wing & Gould, 1979). Kanner's definition of autism was not wrong; it was merely too narrow.

An important question to ask is whether autistic children with normal IQs indeed have the same disorder as autistic children with IQs in the retarded range. Bartak and Rutter (1976) compared autistic boys with nonverbal IQs of 70 or more with autistic boys whose nonverbal IQs were 69 or less. On the phenomena most characteristic of autism, (i.e., social relationships, language, and ritualistic behaviors, they found the two groups essentially alike. All showed serious impairment in the development of social relationships, all were late in talking and showed marked impairment in imaginative play, and all showed some type of stereotyped, ritualistic, or compulsive behavior. The differences between the normal and low IQ groups tended to occur in associated features. For example, self-injurious behaviors and hand or finger stereotypies were much more frequent in the mentally retarded group. The mentally retarded autistic boys were more likely to display extremely deviant social responses such as smelling people. Also, more of the mentally retarded group had been delayed in their motor development.

Even though the two groups are similar on essential features, the prognosis for high functioning autistic
children is much better than for the severely retarded group. This has been borne out by all of the follow-up studies which measured IQ (Campbell, Hardesty, Breuer, & Polevoy, 1978; DeMyer et al., 1973; Lotter, 1974a; Rutter, Greenfeld, & Lockyer, 1967). The higher functioning autistic adolescent or adult has a better chance of being employed in a sheltered setting, and is less likely to require institutionalization. In addition, mentally retarded autistic children are much more likely to develop seizures during adolescence (DeMyer et al., 1973; Rutter, 1970).

The implications to be drawn from the mental retardation research is that it is not sufficient to describe samples of children in terms of whether or not they meet certain autistic criteria. They must be described in terms of their intellectual abilities as well, because IQ is an important factor in the pattern of symptoms which will be displayed and in the prognosis. Furthermore, careful consideration must be given to test selection. Nonverbal tests should be included, and the level of testing must be appropriate to the child's abilities. In some cases that may mean that an infant scale, such as the Bayley Scales of Infant Development (Bayley, 1969), will be used with an older child whose mental age is below 30 months. While appropriate caution is in order, such use is sometimes necessitated by the lack of tests designed for severely
handicapped children, and, more importantly, by the fact that pertinent and meaningful data can be obtained (Chase, 1985; Gerken, 1983).

**Summary**

Confusion about the definition of the syndrome of autism was discussed due to its impact on the interpretation of follow-up studies. The confusion was traced to specific areas where empirical data was challenging original assumptions. The early literature assumed that autism was the earliest form of schizophrenia and autistic children were studied together with others who were psychotic, making their conclusions difficult to interpret. Important distinctions were found between the two disorders, and consequently the definition of autism required reference to age of onset and the absence of hallucinations.

Second, despite their normal physical appearance and their sometimes amazing peak skills, most autistic children are mentally retarded. "Peak skills occur in a small percentage of autistic children and are not a primary feature in a disorder that coexists with mental retardation" (Schopler, 1983, p. 104).

The third edition of the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders* (DSM-III) includes infantile autism under a larger category of Pervasive Developmental Disorders (see Appendix B for the diagnostic criteria). Autism now has the most
specific defining criteria of any of the severe disorders of childhood (Achenbach, 1982), and diagnosis can be made more reliably than ever before (Schopler, 1983).

The struggle with the definition of autism nevertheless continues. Achenbach (1982) points out the difficulty in ruling out hallucinations and other thought disorders in a population of children whose communication skills are severely limited. Tanguay (1984) cites symptom overlap with other syndromes and multiple etiology as evidence for lack of descriptive and etiological validity, respectively. Waterhouse and Fein (1984) report that systems for differential diagnosis of autism and childhood schizophrenia are still inadequate, and that differential diagnosis within the pervasive developmental disorders is not sufficiently reliable and valid (Fein, Waterhouse, Lucci, & Snyder, 1985). Creak observed in 1963 that in attempting differential diagnosis there always seem to be some children who do not fall clearly into our prescribed categories. It is still a valid observation today.

Problems With Objectivity of Measurement

Follow-up studies have also been hindered by a lack of objective measurement, both initially and at follow-up. Initial measures taken have generally included some estimation of the severity of symptoms as well as estimations of intelligence. Severity has been interpreted differently by each researcher, and the problems in the area
of measurement of intelligence reflected the current thinking regarding the testability of autistic children. Follow-up measures have generally attempted to estimate outcome for these children, and such measures have not been characterized by objectivity.

**Measurement of Severity**

No one child shows all the essential and associated features of autism and it is recognized that some children show a more severe form of the syndrome than others. From the beginning, follow-up studies of autistic children recognized that severity of the syndrome was an important factor to consider.

Even as early as Kanner it had been speculated that autism could occur in different degrees of severity and that severity could be a factor in outcome. For Kanner, the presence or absence of speech at the age 5 was an index of the "severity of autistic isolation" (1973, p. 208), and it was indeed a strong predictor of outcome. Of 42 autistic children followed into adolescence, 19 had not talked by the age of 5. All but one of the 19 remained, at follow-up, "in a state of complete isolation and, on superficial observation, can hardly be distinguished from markedly feebleminded persons" (Kanner & Eisenberg, 1955, p. 238). Subsequently, other follow-up studies also attempted to rate severity and correlate it with outcome.

Rutter and his colleagues (Lockyer & Rutter, 1969,
1970; Rutter, Greenfeld, & Lockyer, 1967; Rutter & Lockyer, 1967) reported follow-up data on 63 children diagnosed with infantile psychosis. Later, 56 of these children were identified as "truly autistic" (Rutter, 1970, p. 437), based on onset before 30 months of age and the presence of ritualistic behaviors. Severity of the disorder was found to be predictive of outcome. Each child was rated on a 5-point scale on a list of behaviors. The symptoms included those commonly believed to be associated with autism as well as other problem behaviors associated with non-psychotic children (e.g., aggression and eating disorders). While none of the behaviors significantly discriminated between the autistic group and a matched group of non-psychotic children, a total symptom score, obtained by summing the ratings, was found to be predictive of outcome for the autistic group. The children who were judged to have made a good social adjustment at follow-up had obtained significantly lower symptom scores than the ones who had made only a fair adjustment (Rutter, Greenfeld, & Lockyer, 1967).

DeMyer and her colleagues (DeMyer et al., 1973) reported follow-up information on 120 autistic children, 94 of whom were seen individually at follow-up. The autistic children were classified into three gross categories of high, middle, and low autistic, based upon use of communicative speech and whether or not their intellectual
and perceptual-motor performance abilities were globally retarded or showed inconsistencies. Despite the gross and subjective nature of the classification system, severity of autism was found to be a good predictor of outcome. The high autistic group had the best outcome. Only 25% of this group was autistically withdrawn at follow-up, while 80 percent in the middle and low autistic groups remained withdrawn. Children who were functioning at age level in school when re-evaluated were more likely to have been in the high autistic group.

Thus the severity of the autistic symptoms has been of interest to researchers from the beginning, and different ways of defining severity have proved to be prognostic indicators of outcome, with the less severely affected children having the better chances for an adequate adjustment. Research dealing with autistic children should precisely indicate the extent to which the children show the various characteristics, but this has rarely been done (Balow, 1981). Severity is an important factor which warrants an objective method of measurement.

Objective measurement of the symptoms of autism has been a major area of research (Clancy, Dugdale, & Rendle-Short, 1969; Krug, Arick, & Almond, 1981; Ruttenberg, Dratman, Fraknoi, & Wenar, 1966; Schopler, Reichler, DeVellis, & Daly, 1980). The Childhood Autism Rating Scale (CARS) (Schopler, Reichler, DeVellis, & Daly, 1980) is a 15-
item scale which assesses not only the presence of autism but also the severity (see Appendix C). Each of the 15 items relating to autism is scored on a scale from 1 (no abnormality) to 4 (severe abnormality), with half-point scores possible. In a sample of 537 children, the authors of the scale found that a score of less than 30 identified children who were not autistic. Among the autistic children, scores fell in a bimodal distribution. A score of 36 with a rating of 3 or more on at least 5 scales was designated as severe autism. Scores between the two extremes indicated mild to moderate autism. Use of a rating scale such as the CARS could provide the information on the extent of the various characteristics shown by autistic children which has been lacking in previous studies.

Measurement of Intelligence

Problems associated with the measurement of intelligence in autistic children reflected the current thinking regarding the intellectual potential and thus the testability of autistic children.

In the DeMyer study the method of measuring the children's intelligence changed over the 7-year course of the study. During the first 4 years many of the children were considered untestable and the Vineland Social Maturity Scale was used as a measure of performance IQ. For verbal IQ a verbal mental age was estimated using extensive descriptions of the child’s language which were available
from parent interviews and taped psychiatric interviews. General IQ was estimated by averaging the performance and verbal scores. Later, verbal and nonverbal items from standardized tests were used for estimates of verbal and performance IQ, and general IQ was measured by such tests as the Stanford-Binet, Merrill-Palmer, and Wechsler Intelligence Scale for Children. Despite the obviously inadequate assessment of IQ, the scores were predictive of follow-up status.

The method of obtaining IQ scores in the Lotter follow-up (1966, 1967, 1974a, 1974b) was poorly described, but apparently was based in some subjective manner on a combination of the Peabody Picture Vocabulary Scale, the Seguin formboard, and the Draw-a-man test. When the child proved untestable, educational and hospital records were searched for recent test scores. Four of 32 autistic children had never been successfully tested.

The procedures described in these follow-up studies are clearly inadequate for measuring intelligence. Now that we are aware that autistic children can be functioning at many different levels, a variety of tests are used to measure cognitive ability (cf. Ando, Yoshimura, & Wakabayashi, 1980; Freeman et al., 1981; Schopler, Reichler, DeVellis, & Daly, 1980). As highlighted previously, the emphasis is on measures of nonverbal intelligence due to the basic language deficit of these children and on testing at the appropriate
level of difficulty.

**Measurement of Outcome**

As noted above, outcome measures have not been characterized by objectivity. Different authors have interpreted outcome differently, and placement of subjects into different categories of outcome has necessarily involved some amount of subjective judgment.

Eisenberg (1956) reported follow-up data on 63 autistic children meeting Kanner's definition of the syndrome. The author rated outcome in three categories: "poor," "fair," and "good." A "good" outcome meant that the person was functioning well in academic, social and community levels, and was accepted by peers, even though he or she might still be considered somewhat odd. A "fair" outcome meant that the person was able to attend regular classes at appropriate age level and had some meaningful contacts with other people, even though exhibiting "schizoid peculiarities" (1956, p. 608). A "poor" outcome indicated severely maladaptive current functioning due to apparent severe retardation or grossly disturbed behavior. Of the 63 children, 3 had achieved a "good" outcome, 14 a "fair" one, and 46 a "poor" one.

Eisenberg found the presence of useful speech by the age of 5 to be an important predictor of the outcome. Thirty of 31 children who were not talking by that age had a "poor" outcome, whereas 16 of 32 speaking children had a "poor" outcome. In retrospect, it is obvious that a huge gap
existed between Eisenberg's "fair" and "poor" outcomes.

In the Rutter study outcome was determined by a social adjustment rating of "good," "fair," "poor," or "very poor" (Rutter, Greenfeld, & Lockyer, 1967). A rating of "good" indicated a normal or near-normal social life and satisfactory functioning in school or work. "Fair" indicated social and educational progress despite significant abnormalities in behavior or interpersonal relationships. "Poor" indicated a severely handicapped individual unable to lead an independent life, but who, it was felt, had some potential for social progress. "Very poor" indicated total inability to lead an independent life and no such estimated potential. At follow-up social adjustment was rated "good" for 14% of the subjects, "fair" for 25%, "poor" for 13%, and "very poor" for 48%. Four factors were found to be correlated with outcome: initial IQ, speech, severity, and the amount of schooling the child had had.

Lotter (1974b) used the same social adjustment ratings as Rutter, Greenfeld, and Lockyer (1967), described above. The findings were remarkably similar, including the same predictors. The Social Quotient from the Vineland Social Maturity Scale was also a strong predictor of outcome.

In the DeMyer study a rating of the work/school functioning of the autistic child was chosen as an outcome measure. The scale had 5 possible ratings which took into
account the child's age, preschool and school-age levels for initial evaluation and high school and post-high school age levels for follow-up. The range, for example, for post-high school age was from class I (Normal), in college or technical training or holding a job without difficulties, to class V (Sub-trainable), cannot work or do any household tasks, even with supervision, must be taken by another person for excursions into the community.

As mentioned earlier, DeMyer divided the autistic children into three categories of high, middle, and low. At follow-up 71% of the low autistic group were in Class V (sub-trainable), 48% of the middle group, and 5% of the high group. The authors found that the best prediction of the child's work/school functioning at follow-up was the work/school rating that had been made at initial evaluation. Additional predictors were IQ, severity of the disorder, the ratings on a social skills scale and an conversational speech scale, and an index of brain dysfunction.

In conclusion, there appears to be no recognized standard for the evaluation of outcome. With the exception of Lotter, each researcher devised his or her own rating scale, and even some recent follow-up studies have reported outcome mainly in the form of case study information (e.g., Campbell, Hardesty, Breuer, & Polevoy, 1978; Shell, Campion, Minton, Caplan, & Campbell, 1984).

Of the outcome measures reviewed, the one used by
DeMyer et al. (1973) appears to have several advantages. The work/school functioning scale taps adaptive functioning and takes into account the age of the child, adolescent, or adult. Classifications within its five categories are described in terms of the school or job-related work a person is doing, a less subjective decision than judging, for example, the potential for social progress.

Lotter (1978) also cited some advantages of the work/school scale: its simplicity, prediction of practical information, usefulness with preschool children, and ease of interpretation with little correction for changes in age. Thus in the absence of an accepted standard procedure for the evaluation of outcome, the work/school scale would appear to meet some of the basic criteria one would select for such a measurement.

Summary

The difficulties researchers have had with objective measurement of autistic children were reviewed. Initial evaluations of this population generally included some estimate of the severity of the autistic symptoms and of the level of cognitive functioning. Since Kanner, severity had been recognized as an important factor, but no consensus was ever reached about how to measure it. Each researcher devised a unique system. The merits of newer and more objective methods of assessing severity, such as the Childhood Autism Rating Scale, were discussed.
The problems associated with measuring intelligence in this population reflected the current thinking regarding the intellectual potential and testability of the children. Intelligence was measured by all kinds of nonstandard methods, was often subjectively estimated, and sometimes poorly described in the study.

Last, the measurement of outcome in the follow-up evaluations was also characterized by a lack of objectivity and by a variety rather than a consensus of methods. The work/school functioning scale devised by DeMyer et al. (1973) was highlighted as a comprehensive and objective measure for evaluating outcome.

Subclinical Autism: Children Who Are Autistic-Like

In the discussion of autism and the particular problems related to follow-up studies of autistic children, it has been noted that autism occurs in varying degrees of severity. The focus has thus far been restricted to the autistic end of the spectrum. There is apparently a large number of young children who show one or more autistic symptoms, but do not meet all the criteria for diagnosis. Evidence for this comes from several studies that have looked at populations other than autistic.

Wing and her associate (Wing, 1978, 1981; Wing & Gould, 1979) have reported an epidemiological study designed to identify all children in a particular geographic area of England who, regardless of their diagnosis, showed one or
more of the following: (a) severe impairment of social interaction, (b) language abnormalities, and (c) repetitive stereotyped behaviors. Without taking age of onset into consideration, they found a prevalence rate of 21 per 10,000, and there was a marked tendency for the three problems to occur together. All socially impaired children had repetitive stereotyped behaviors, and almost all had language abnormalities as well. In line with current reported prevalence rates for autism, a subgroup representing an incidence of 5 per 10,000 could be clearly identified as autistic. Undoubtedly Wing's larger group included many schizophrenic children and other children with severe language disorders, who, according to Rutter (1978), are different from autistic children in important ways. It does, however, illustrate the point that the population of children who show autistic behaviors is much wider than the subgroup who can be diagnosed autistic.

Interestingly, in the research that led Rutter to state that children with severe developmental receptive language disorders (dysphasics) were clearly different from autistic children (Bartak, Rutter, & Cox, 1977), a sub-group of children who clinically could not be diagnosed autistic or dysphasic was termed "mixed" and indeed appeared quite mixed on their test results. Individual cases scored like the autistic group on some measures and like the dysphasic group on others. Rutter (1979) summarized distinct differences
between autistic children and dysphasic children, including their understanding of language, pattern of subtest scores on the WISC, use of imaginative play and understanding of gesture, use of speech for social communication, and abnormal use of language other than simple echoing. He did not diagnose or attempt to explain the so-called "mixed" group which made up 11% of his sample of 47 children.

Paul and Cohen (1984) reported outcomes of severe disorders of language acquisition. One-half of their original group of 28 children had some autistic symptoms, such as poor social relating and inappropriate play, but of insufficient severity to diagnose as autistic. The authors reasoned that the children probably could be diagnosed as Atypical Pervasive Developmental Disorder, a new diagnostic category in the DSM-III, but they questioned the usefulness of a diagnostic system where so many children fall through the cracks and need to be placed in a "grab-bag category" (Paul & Cohen, 1984, p. 419). The authors referred to an "autistic end of the spectrum of language disorders" (Paul, Cohen, & Caparulo, 1983, p. 527).

Fisher, Burd, and Kerbeshian (1987) propose that Atypical Pervasive Developmental Disorder (APDD) and autism are the same conditions, differing only in severity of symptoms in social relatedness and language impairment. They found that the autistic children in their state-wide survey did not differ from APDD children on intelligence
test scores, but did score significantly lower on measures of expressive and receptive language. In addition, they found a significant negative correlation between age and the autistic group's scores on intelligence and language tests, suggesting that the rate of improvement for autistic children slows relative to age; this was not found for the APDD children.

More data and better specification of the subclinical autistic-like conditions is badly needed (Rutter & Schopler, 1987; Volkmar, 1987). We can speculate that they represent milder forms of the same disorder, but we do not have enough evidence to be convincing. Of course, differentiation of autism from such syndromes could be unequivocal only if we had identified a unique and autism-specific set of diagnostic criteria, and we have yet to do that (Rutter & Schopler, 1987).

**Summary**

Autism is a lifelong handicapping disorder, and the changes which can occur with development make follow-up of individuals and groups of autistic children essential. Interpretation of available follow-up studies has been hindered by significant problems. First, the struggle with the definition of the syndrome were discussed. Definitions among studies varied due to assumptions about autism's relationship to schizophrenia. In addition, studies were hindered by inadequate and inappropriate assessment of
intelligence due to assumptions about the average intellectual potential of autistic children. Eventually research led to clearer statements of the definition of autism. These statements specified age of onset and the absence of hallucinations, and acknowledged the presence of different levels of intellectual functioning. The DSM-III criteria is an example (see Appendix B), and represents the definition of autism used in this study.

Second, the problems with objectivity in measurement were discussed. Early researchers knew that intelligence, severity, and outcome were all important factors to measure, but accurately measuring each of them proved difficult. It took many years for the intellectual assessment of autistic children to become adequate and appropriate. It involved a recognition that "untestable" meant a lower level of cognitive ability and that the basic language deficit of these children called for the use of nonverbal tests. In the areas of severity and outcome, the problem appears to have been mainly a lack of an accepted standard for the measurement of these essential factors. Despite the imprecise measurement of autistic children, it should be noted that follow-up studies nevertheless found IQ and severity both to be predictive of outcome.

Last, autistic-like children were discussed. Autistic symptoms occur in children who are not considered to be autistic. This subclinical group represents many unknowns
in the study of severe childhood psychopathology. They have not been studied per se, but have shown up in studies of children with severe language deficits, and in numbers significant enough to warrant more thorough investigation. If autism occurs on a continuum, the subclinical end of that continuum should be of as much interest as the other extreme.

Purpose of the Study

Several important changes have occurred since the follow-up studies cited above, and more recent follow-up studies are lacking. First, as noted above, the defining criteria for infantile autism are more clearly delineated than ever before. The DSM-III adopted the essential characteristics for which empirical support had been found (Rutter, 1978). Second, objective measurement of intelligence is no longer hindered by preconceived notions about average intellectual potential. Third, objective measurement of the symptoms of autism has been a major area of research, resulting in instruments such as the CARS that enable us to describe both the presence and the severity of autism. Research on autism has led to specific improvements in our ability to identify autistic children and assess their functioning. Early researchers had to struggle with diagnostic and measurement problems which are not nearly so formidable today.

The changes and refinements which have occurred warrant
a replication of follow-up studies with the autistic population today. In addition, the lack of information regarding the characteristics and prognosis for a larger group of children called autistic-like represents an important area of research that has been neglected. Questions about both populations of children today are answerable by research methods, and are of direct concern to clinicians working with autistic and autistic-like children, their parents, and their teachers.

Subjects in the present study were selected based upon the presence of autistic characteristics. The number and severity of autistic symptoms determined placement in one of three categories at follow-up: nonautistic, mildly/moderately autistic, and severely autistic. Specific DSM-III criteria were used for diagnosis, IQ was measured with standardized and appropriate verbal and nonverbal tests, and severity and outcome were measured with objective and previously-applied instruments. Of particular interest in the present study is the question of how children who fall at different points along an autism continuum are alike and how they are different in terms of initial presentation and developmental prognosis.
CHAPTER II

METHOD

Subjects

Thirty-five children were seen for follow-up evaluation. Twenty-eight were children who had originally been evaluated at an out-patient child clinic and whose records indicated autistic or autistic-like behaviors. These children had originally qualified for a general diagnosis of Pervasive Developmental Disorder in that they showed multiple delays in language and social development. In addition, they had demonstrated an onset of symptoms prior to 30 months of age.

The clinic is a United Way Agency whose fees are determined by a sliding scale depending upon income. It is the only such child agency in a large metropolitan area and regional schools routinely refer severely handicapped children to the facility. Refer to Table 1 for a demographic description of the 28 clinic subjects. Minimum length of time for follow-up was 2 years. The average length of time that elapsed between initial and follow-up evaluations was 4 years, 6 months (SD = 36.9 months, range 2 years to 15 years, 8 months).

Seven diagnosed autistic children from local school
Table 1

Subject Characteristics

<table>
<thead>
<tr>
<th></th>
<th>Clinic</th>
<th>School</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Initial age (years-months)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>1-3 to 5-5</td>
<td>3-8 to 12-2</td>
</tr>
<tr>
<td>Mean</td>
<td>3-2</td>
<td>6-9</td>
</tr>
<tr>
<td><strong>Age at follow-up (years-months)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>3-5 to 18-3</td>
<td>7-2 to 15-11</td>
</tr>
<tr>
<td>Mean</td>
<td>7-8</td>
<td>11-6</td>
</tr>
<tr>
<td><strong>Ethnicity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>20</td>
<td>2</td>
</tr>
<tr>
<td>Black</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Hispanic</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Asian</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>24</td>
<td>7</td>
</tr>
<tr>
<td>Female</td>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>

districts were added in order to increase the number of subjects in the study. All of the children from the local schools had been identified as autistic by a diagnostic team which always consisted of a psychologist, a speech pathologist, and an educational diagnostician, as per state guidelines. Table 1 also presents descriptive data on the
seven autistic school children. The average length of time between initial and follow-up evaluations for the school children was 4 years, 8 months \((SD = 23.3\) months, range 2 years, 7 months, to 8 years, 7 months.

Thus 7 diagnosed autistic children were added to the original subject group of 28, resulting in initial and follow-up data on 35 children. All 35 children were seen individually at follow-up.

Materials

Childhood Autism Rating Scale (CARS)

The CARS (Schopler, Reichler, DeVellis, & Daly, 1980) (Appendix C) is a 15-item scale which assesses the presence and severity of autism. Each of the 15 items relating to autism is scored on a scale from 1 (no abnormality) to 4 (severe abnormality), with half-point scores possible. In a sample of 537 children, the authors of the scale found that a score of less than 30 identified children who were not autistic. Among the autistic children, scores fell in a bimodal distribution. A score of 36 with a rating of 3 or more on at least five scales was designated as severe autism. Scores between the two extremes indicated mild to moderate autism. The authors reported that interrater reliability between two trained raters on 280 cases averaged .71, with reliability for individual scale items ranging from .55 (Inconsistencies in intelligence) to .93 (Relationships with people). Coefficient alpha was .94,
indicating high internal consistency (Parks, 1983). Validity was assessed by comparing scale scores and clinicians' ratings of autism obtained at the same evaluation session ($r = .84$, $p < .001$), and by comparing total scores with independent clinical assessments made by a child psychiatrist and a child psychologist ($r = .80$, $p < .001$). In addition, Teal and Wiebe (1986) compared three autism assessment instruments on their ability to discriminate between autistic children and matched trainable mentally retarded children. The CARS predicted group membership best, with 100% accuracy and no overlap of groups. These reliability and validity results have been replicated (McCallon & Garfin, 1986; Rebiliet, 1986).

CARS raters are trained by viewing video tapes provided by the authors of the scale, accompanying trained raters as they observe children, and obtaining a reliability of at least .80 with the trainers. The author was trained by this procedure for a previous study (McCallon, 1986) and made all the CARS ratings for subjects in this study.

**Autism Checklist**

A 14-item checklist (Clancy, Dugdale, & Rendle-Short, 1969) (Appendix D) was used to assess the presence of autistic symptoms at the time of initial evaluation. This checklist is used as a screening measure by the National Society for Autistic Children and Adults. In devising the scale, the authors analyzed the symptoms of a group of 25
children who had been clinically diagnosed as having infantile autism according to the Creak criteria plus onset in the first three years of life. The symptoms of these autistic children were compared with those of normal, retarded, rubella deaf, and cerebral palsied children. Fourteen core symptoms were identified which differentiated the autistic group, and the authors recommended that 7 of the 14 behaviors should be present for a diagnosis of autism to be considered. Checklists are not intended for diagnosis (Freeman & Ritvo, 1981), but the 14 items have been found to be a core of symptoms which can describe the behavior of autistic children and adolescents, and which do not describe the behavior of nonautistic ones (McCallon, 1986). Reliability data for the Autism Checklist are not available.

For the process of searching through initial medical and psychological or school evaluations in order to identify the presence of autistic symptoms, a checklist was deemed more appropriate than the CARS which would have required inferential ratings of severity.

Problem Checklist and Parent Questionnaire

During the initial clinic evaluation, parents completed a 35-item Problem Checklist, information on developmental milestones, and a Parent Questionnaire on family functioning, sibling relationships, and discipline. The Problem Checklist (Appendix B) covered a wide range of child behavior difficulties which are not necessarily
representative of autism. The Parent Questionnaire is shown in Appendix F. At follow-up, parents filled out another Problem Checklist, based on current behavior, and another Parent Questionnaire.

**Vineland Social Maturity Scale**

The Vineland, a measure of adaptive functioning (Doll, 1965) was scored for each child at follow-up based on parent interview information. The 1965 edition, rather than the revised edition (Sparrow, Balla & Cicchetti, 1984) was used so that Social Quotient (SQ) scores would be comparable with those collected on the children initially. The Social quotient is computed by a ratio method (Social Age/Chronological Age X 100). The mean score is 100; unfortunately a major drawback is that the standard deviations of the SQ are not constant with age. Nonetheless, useful information about adaptive functioning is gained (Sattler, 1982).

**Tests of Intelligence/Developmental Functioning**

Measures of cognitive functioning were taken from the report of the child's initial psychological evaluation at the clinic or school. When possible, due to the basic language deficit of this population, a nonverbal measure of functioning was chosen over a verbal one for comparison. Examples of essentially nonverbal tests include the Bayley Scales of Infant Development (Bayley, 1969), the Leiter International Performance Scales (Leiter, 1948), the
Merrill-Palmer Scale of Mental Tests (Stutsman, 1931), and the Performance scales from the McCarthy Scales of Children's Abilities (McCarthy, 1972) and the Wechsler Intelligence Scale for Children--Revised (Wechsler, 1974).

At follow-up a nonverbal measure of cognitive functioning was obtained for each child. The tests utilized were the Bayley Scales of Infant Development, the Leiter International Performance Scales, and the Performance Scale of the Wechsler Intelligence Scale for Children--Revised.

Language Assessment

The Peabody Picture Vocabulary Test--Revised (PPVT-R) (Dunn & Dunn, 1981), was utilized as a language measure. The PPVT-R is a standardized test of receptive vocabulary, with a mean of 100 and standard deviation of 15. Standard scores range from 40 to 160.

Outcome Rating

A work/school rating scale was used as an outcome measure (Appendix G). It was derived from one used by DeMyer et al. (1973), with the only changes being related to the fact that all autistic children are now entitled to public education, a change from the time DeMyer's work/school rating was devised. For example, some categories originally included such problems as "not recommended for high school" (1973, p. 208), which is no longer relevant.

Procedure

Forty-seven children who had been evaluated at the out-
patient child clinic and whose medical records indicated autistic or autistic-like behaviors were selected for follow-up. A review of their records verified that they initially qualified for a general diagnosis of Pervasive Developmental Disorder, in that they showed multiple delays in social and language development. A minimum length of follow-up time of two years was required, according to the method employed by DeMyer et al., 1973. Forty-one children met this criteria. Letters were sent to the last known addresses of the clients explaining the research study and stating that they would be contacted by phone to see if they were interested in participating (Appendix H). Five families had moved away with no forwarding address; no current telephone number could be found for an additional four families who also failed to respond to a second letter. Of the 32 families contacted, 2 were not interested in participating. Thus 30 children were seen. Subsequently 2 of these were eliminated from the study because age of onset before 30 months could not be verified in either the parent interview or the child's records. Of the original 47 clinic children, 28 subjects who initially showed a pervasive developmental delay in language and social skills with onset prior to 30 months of age were included in the follow-up study.

Six autistic school children were added through the help of a local school district's autism program. The
special education department of the district sent letters to 20 parents of autistic children (Appendix I) recommending participation in the research study and giving a telephone number for the parents to call if they were interested in bringing their child to the clinic for a free evaluation. Eight parents called, for a response rate of 40%; subsequently six children were evaluated. One of the children whose parent called was a client of the clinic and had already been included in the study; another parent failed to keep two appointments that were made.

One other autistic child was added to the study. This child lived in nearby city, and initial and follow-up data were available due to the author having done an evaluation on the child for the school district.

Signed parental permission was obtained on all children tested. See Appendix J for the permission form.

Collection of Initial Data

Before any children were seen, the files of 47 clinic children were examined to verify that the child did indeed meet the general criteria for a Pervasive Developmental Disorder and that a minimum requirement of 2 years follow-up time would be met. Reports were searched for references to a delay in social and language development. Information that would be needed for the study was coded at that time except that, in order to avoid influencing the author, no note was made of the diagnosis the child received. It
should be noted that most of the children's records span several years of test results, and diagnoses frequently changed over time. Often the initial evaluation conservatively diagnosed only a developmental delay and it was recommended that the child attend a special diagnostic school run by the clinic in order to clarify diagnostic issues. Information on 41 clinic children was coded. The author did not examine the child's file again until after the child was seen and all follow-up measures scored. The time between coding information from the files and seeing the child was a minimum of 2 months and in most cases (17 of 28) was over one year.

Records for the school sample children were not searched for initial data until after the children had been evaluated and all measures scored. Generally the most complete evaluation information was available from the assessment done to qualify the student for inclusion in the autism program. Length of follow-up time was counted from that evaluation.

Initial evaluation information recorded for each subject included the age of the child at initial evaluation, sex, ethnicity, IQ, Vineland score (Social Quotient), PPVT or PPVT-R, Autism Checklist scores, and, for the clinic children, the Problem Checklist, Parent Questionnaire, and developmental milestones information.

IQ scores were available on 27 children, the other 8
being described in the report as "untestable." For five subjects, no nonverbal measure had been given and their score from the Stanford-Binet Intelligence Scale was used. The tests utilized are shown in Table 2.

Table 2

<table>
<thead>
<tr>
<th>Test Utilized</th>
<th>No. Tested</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bayley Scales of Infant Development</td>
<td>14</td>
</tr>
<tr>
<td>McCarthy Scales of Children's Abilities</td>
<td></td>
</tr>
<tr>
<td>Perceptual-Performance Scale</td>
<td>5</td>
</tr>
<tr>
<td>Merrill-Palmer Scale of Mental Tests</td>
<td>2</td>
</tr>
<tr>
<td>Stanford-Binet</td>
<td>5</td>
</tr>
<tr>
<td>Leiter International Performance Scales</td>
<td>1</td>
</tr>
</tbody>
</table>

Initial SQ scores were available for 22 of the 35 children. Only 9 children had been given a Peabody Picture Vocabulary Test in their initial evaluations; 7 had been given the original test, and 2 the revised edition. When scores from the PPVT were the only ones available, a table in the PPVT-R manual (Dunn & Dunn, 1981) was used to convert the standard score to one comparable to the PPVT-R.

The Autism Checklist was scored based on information contained in the child's initial medical and psychological or school reports. If a particular behavior was described
in any of the evaluations made at initial referral, the item was checked as present.

The initial information on six developmental milestones was used to make a global rating of motor development delay. Each of the items was judged to represent normal or delayed motor development, using the norms from the motor scale of the Bayley Scales of Infant Development (Bayley, 1969) (see Appendix K). According to Sattler (1982), the Bayley is the best available measure of infant development. The global rating consisted of the number of items on the six-item scale which were delayed (at the 95th percentile or above).

Collection of Follow-Up Data

At follow-up, children were given a language and cognitive evaluation using standardized tests. The parents were asked to fill out a current Problem Checklist and Parent Questionnaire. The parents were interviewed in order to determine age of onset, current behavior, present school placement, and adaptive functioning level. Age of onset was defined as evidence of developmental impairment or distortion, as recommended by Rutter and Schopler (1987). Thus, if a parent reported that the child was not cuddly as a baby or did not like being held, age of onset was arbitrarily recorded as 6 months, and if a child of 3 years or more had never talked, age of onset was recorded as 24 months.
Based upon the observation of the child in the testing situation and parent interview information regarding current behavior, the author scored the Autism Checklist and the CARS, rated the child on the Work/School scale, and made a diagnosis, for research purposes only, using DSM-III criteria. Ten of the Vinelands were scored by another graduate psychology student who interviewed the parents at the time the child was being tested.

A nonverbal measure of cognitive functioning was obtained for each child. The tests utilized were the Bayley Scales of Infant Development ($n = 5$), the Leiter International Performance Scales ($n = 28$), and the Performance Scale of the Wechsler Intelligence Scale for Children—Revised ($n = 2$). The five follow-up cases in which the Bayley was used all exceeded the age norms for the test, therefore an IQ was computed by finding the mental age at which the obtained score was equal to an average IQ of 100, then dividing that mental age by the child's chronological age and multiplying by 100. The Leiter was administered according to the original author's instructions (Leiter, 1948); IQ's for this test are computed by the ratio method and are adjusted by adding 5 points to the derived score. A table provided by Levine (1984) was utilized for making this computation.

At follow-up, 26 of the 35 children were able to take the Peabody Picture Vocabulary Test—Revised. Seven
children were able to take the test, but their scores fell below the norms for the test. For these cases only, the child's obtained age equivalent on the test was divided by his or her chronological age and multiplied by 100 to provide a score that could be used in computations.

Reliability of Diagnosis

As the study was originally designed, the children were to be seen by the medical staff at the clinic after their psychological evaluation, providing a second opinion of current diagnostic status. Nine children were seen by a pediatrician who knew that a follow-up of autistic children was being done. When Autism and Residual Autism are assumed to represent the same syndrome, which is consistent with the developmental literature and the latest diagnostic guidelines (American Psychiatric Association, 1987), agreement between the author and the pediatrician on current diagnosis was 67%. Table 3 describes the three discrepancies in diagnosis. Two of the three discrepancies involved the mis-diagnosis of Residual Autism in a pair of twins. These children had never been diagnosed as autistic in the four years that they had been followed at the clinic, therefore the pediatrician's diagnosis of Residual Autism, which requires that the child once met the full criteria for Infantile Autism, was inappropriate.

Six children who were to be seen by the pediatrician never kept their appointments and so records were searched
Table 3

Discrepancies in Diagnosis

<table>
<thead>
<tr>
<th>Case</th>
<th>Author's Diagnosis</th>
<th>Physician's Diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>#33</td>
<td>Developmental Language Disorder</td>
<td>Residual Autism</td>
</tr>
<tr>
<td>#15</td>
<td>Developmental Language Disorder</td>
<td>Residual Autism</td>
</tr>
<tr>
<td>#16</td>
<td>Atypical Pervasive Developmental Disorder</td>
<td>Possible Residual Autism</td>
</tr>
</tbody>
</table>

for a recent diagnosis. Three of the six had had full multi-disciplinary team evaluations at the clinic within 4 months of their participation in the research project. Another had had a multi-disciplinary evaluation in another city just 2 months prior to follow-up and the parents gave permission for the records to be forwarded to the clinic. For the fifth child a multi-disciplinary team at the child's school had diagnosed him autistic and permission was obtained to use his school records in the research. Agreement of diagnosis between the author and any of these multi-disciplinary teams (n = 5) was 100%. This left one child for whom a recent concurrent diagnosis could not be found. He had been consistently diagnosed autistic at the clinic, and in the author's opinion, continued to meet the criteria for that diagnosis.

After the first 15 children had been seen at the clinic
it was no longer feasible for the physicians to see the children in the follow-up study. Subsequently, audio tape recordings were made of the parent interview \((n = 13)\). This tape recording and a behavioral description of the child's test behavior was given to one of two licensed psychologists who specialize in working with children, and they provided a second diagnostic opinion. These psychologists knew that the research was in general related to autism, but did not know the specifics of the study or how the children were selected. Any mention of diagnosis on the tape was erased, as well as one mention of the use of the drug Fenfluramine because of its association with the treatment of autism. Reliability of agreement was 92%. The one disagreement involved a child the author diagnosed as autistic and the psychologist diagnosed mentally retarded "with autistic features."

The seven autistic children from the school sample had all been diagnosed by multi-disciplinary teams, and at the time of follow-up evaluation they continued to meet the criteria for such a diagnosis. Thus, agreement between the author and any of the multi-disciplinary teams, including the school teams \((n = 12)\), was 100%. Combined with the diagnoses from the pediatricians and the psychologists, overall agreement on diagnosis was 88%. Table 4 provides the distribution of current diagnoses, as given by the author.
Table 4

**Diagnostic Categories--Current**

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infantile Autism or Residual Autism</td>
<td>23</td>
</tr>
<tr>
<td>Developmental Language Disorder</td>
<td>5</td>
</tr>
<tr>
<td>Atypical Pervasive Developmental Disorder</td>
<td>3</td>
</tr>
<tr>
<td>Cerebral palsy</td>
<td>1</td>
</tr>
<tr>
<td>Separation Anxiety Disorder</td>
<td>1</td>
</tr>
<tr>
<td>Schizoid Personality Disorder</td>
<td>1</td>
</tr>
<tr>
<td>Does not currently meet criteria for any disorder</td>
<td>1</td>
</tr>
</tbody>
</table>

Combining Samples

Before combining the 28 clinic children and the 7 school children into one subject sample, the school children were compared with the diagnosed autistic children from the clinic (n = 16) to determine if there were significant differences between the two groups on initial characteristics. T-tests revealed that the two samples of autistic children were not significantly different on NVIQ, SQ or Autism Checklist scores. Table 5 compares the two samples on these measures.

The two samples differed on age at initial evaluation; the school group (M = 6 years, 9 months) was older than the clinic group (M = 3 years), t(21) = 4.86. It seems
Table 5

Comparison of Two Samples of Autistic Children

<table>
<thead>
<tr>
<th></th>
<th>Clinic Children</th>
<th>School Children</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NVIQ</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>51.4 (n = 11)</td>
<td>44.3 (n = 6)</td>
</tr>
<tr>
<td>Range</td>
<td>22 to 94</td>
<td>14 to 74</td>
</tr>
<tr>
<td><strong>SQ</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>64 (n = 8)</td>
<td>53.5 (n = 6)</td>
</tr>
<tr>
<td>Range</td>
<td>45 to 85</td>
<td>26 to 79</td>
</tr>
<tr>
<td><strong>Autism Checklist</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>8.3 (n = 16)</td>
<td>9.4 (n = 7)</td>
</tr>
<tr>
<td>Range</td>
<td>4 to 14</td>
<td>7 to 12</td>
</tr>
</tbody>
</table>

reasonable that that would be the case, since the records taken from school files did not necessarily represent the first time the child had been evaluated for developmental problems. The initial data generally came from the first full evaluation the child had for the autism program in his or her school district. Initial evaluation data for the clinic sample, on the other hand, generally represented the first time the child was evaluated by anyone for developmental problems. With an average age at initial evaluation of 3 years, these children had not yet started school.

Since the two samples of autistic children were not
significantly different on initial intelligence levels, adaptive functioning, and number of autism characteristics evidenced, the 7 school subjects were combined with the larger sample of clinic children to make a subject total for the study of 35 children. The average length of time between initial and follow-up evaluations for the entire sample was 4 years, 6 months, ranging from a minimum of 2 years to a maximum of 15 years, 8 months.
CHAPTER III

RESULTS

Results will be reported as follows: (a) comparisons over time, (b) comparisons between groups based on the severity of autistic symptoms, (c) correlations among measures, and (d) prediction of variables from multilinear regression analysis. Not all data were available on every subject; Ns are noted when data were missing. Due to the number of analyses, an alpha level of .01 was chosen.

Comparisons Over Time

Paired t tests were used to compare changes from early to later evaluations. Initial and follow-up test scores are summarized in Table 6.

Initially, measures of nonverbal intellectual, adaptive, and language functioning, with respective means of 57.6, 65, and 57.3, averaged in the retarded range; however, on all 3 measures, scores spanned the entire range from profoundly or severely retarded to normal. Nonverbal IQ improved significantly over time ($t(26) = -3.43$); adaptive functioning and language scores did not change.

A breakdown of the IQ distribution initially and at follow-up is shown in Table 7. Fifty-nine percent of initially testable children improved enough to change from
one IQ category to another (e.g., from mildly retarded to borderline).

Table 6
Initial and Follow-Up Evaluation Scores

<table>
<thead>
<tr>
<th></th>
<th>Initial</th>
<th>Follow-Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonverbal IQ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>57.6 (n = 27)</td>
<td>68.8</td>
</tr>
<tr>
<td>Range</td>
<td>14 to 110</td>
<td>11 to 120</td>
</tr>
<tr>
<td>Vineland SQ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>65 (n = 22)</td>
<td>65.9</td>
</tr>
<tr>
<td>Range</td>
<td>26 to 90</td>
<td>15 to 126</td>
</tr>
<tr>
<td>Peabody Picture Vocabulary Test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>57.3 (n = 9)</td>
<td>62.5 (n = 27)</td>
</tr>
<tr>
<td>Range</td>
<td>32 to 91</td>
<td>16 to 114</td>
</tr>
<tr>
<td>Autism Checklist</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>8</td>
<td>5.9</td>
</tr>
<tr>
<td>Range</td>
<td>3 to 14</td>
<td>0 to 12</td>
</tr>
<tr>
<td>Problem Checklist</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>8 (n = 23)</td>
<td>8.3</td>
</tr>
<tr>
<td>Range</td>
<td>3 to 16</td>
<td>0 to 20</td>
</tr>
</tbody>
</table>

a Significant change over time.

The number of autistic symptoms displayed by the children, as measured by the Autism Checklist, decreased
Table 7

Distribution of IQ Scores

<table>
<thead>
<tr>
<th>Category</th>
<th>Initial</th>
<th>Follow-Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal 90 and above</td>
<td>2 (7.4%)</td>
<td>9 (25.7%)</td>
</tr>
<tr>
<td>Low Average 80-89</td>
<td>2 (7.4%)</td>
<td>9 (25.7%)</td>
</tr>
<tr>
<td>Borderline 70-79</td>
<td>3 (11.1%)</td>
<td>1 (2.9%)</td>
</tr>
<tr>
<td>Mildly Retarded 50-69</td>
<td>10 (37.0%)</td>
<td>6 (17.1%)</td>
</tr>
<tr>
<td>Moderately Retarded 35-49</td>
<td>6 (22.2%)</td>
<td>4 (11.4%)</td>
</tr>
<tr>
<td>Severely/Profoundly Retarded &lt;35</td>
<td>4 (14.8%)</td>
<td>6 (17.1%)</td>
</tr>
</tbody>
</table>

over time ($t(34) = 3.99$) (refer back to Table 6). Initially children showed a mean number of 8 symptoms out of a possible 14; this changed to 5.9 at follow-up. The Autism Checklist is generally used as a screening instrument, with a score of 7 qualifying a child for further evaluation. Thirteen children on initial evaluation scored below this cutoff; eight had a score of 6, two each had scores of 4 and 5, and one had a score of 3. Table 8 shows the frequency with which individual symptoms on the Autism Checklist were displayed by the children, both initially and at follow-up. The number of symptoms evidenced by the children on the Problem Checklist (out of a possible 35) remained stable over time (refer back to Table 6). Items which were checked by more than half of the parents, either initially or at
Table 8

Frequency of Displayed Symptoms From the Autism Checklist

<table>
<thead>
<tr>
<th>Symptom</th>
<th>Initial</th>
<th>Follow-Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standoffish manner</td>
<td>94%</td>
<td>69%</td>
</tr>
<tr>
<td>Repetitive and sustained odd play</td>
<td>83%</td>
<td>51%</td>
</tr>
<tr>
<td>No eye contact</td>
<td>80%</td>
<td>51%</td>
</tr>
<tr>
<td>Great difficulty in mixing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>with other children</td>
<td>77%</td>
<td>66%</td>
</tr>
<tr>
<td>Acts as deaf</td>
<td>69%</td>
<td>26%</td>
</tr>
<tr>
<td>Prefers to indicate needs by gestures</td>
<td>57%</td>
<td>29%</td>
</tr>
<tr>
<td>Marked physical overactivity</td>
<td>57%</td>
<td>40%</td>
</tr>
<tr>
<td>Not cuddly as a baby</td>
<td>54%</td>
<td>57%</td>
</tr>
<tr>
<td>Resists learning</td>
<td>49%</td>
<td>37%</td>
</tr>
<tr>
<td>Resists change in routine</td>
<td>46%</td>
<td>34%</td>
</tr>
<tr>
<td>Laughing and giggling for no apparent reason</td>
<td>43%</td>
<td>51%</td>
</tr>
<tr>
<td>Lack of fear about realistic dangers</td>
<td>40%</td>
<td>49%</td>
</tr>
<tr>
<td>Spins objects</td>
<td>37%</td>
<td>20%</td>
</tr>
<tr>
<td>Unusual attachment to objects</td>
<td>14%</td>
<td>17%</td>
</tr>
</tbody>
</table>

follow-up, are shown in Table 9.

Some changes appeared in the information gathered from the Parent Questionnaire (Appendix E). Referring to Table 10, parents were less likely to report that their child
Table 9
Most Frequently Checked Problems From the Problem Checklist

<table>
<thead>
<tr>
<th>Problem</th>
<th>Initial</th>
<th>Follow-Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Often does not seem to listen</td>
<td>91%</td>
<td>46%</td>
</tr>
<tr>
<td>Difficult to get his/her attention</td>
<td>87%</td>
<td>30%</td>
</tr>
<tr>
<td>Needs a lot of supervision</td>
<td>52%</td>
<td>52%</td>
</tr>
<tr>
<td>Stubborn</td>
<td>52%</td>
<td>58%</td>
</tr>
<tr>
<td>Easily distracted</td>
<td>44%</td>
<td>73%</td>
</tr>
<tr>
<td>Has difficulty keeping attention on school</td>
<td></td>
<td></td>
</tr>
<tr>
<td>work or other tasks until completed</td>
<td>44%</td>
<td>58%</td>
</tr>
<tr>
<td>Has difficulty waiting turn</td>
<td>13%</td>
<td>52%</td>
</tr>
</tbody>
</table>

seems distant from or disrupts the family. In the matter of discipline, parents were less likely to find spanking effective and more likely to prefer verbal discussion. They were also less likely to report that the child disregards or rebels against the parents' disciplinary measures.

Some data were taken only at initial evaluation. Ten percent of the children had suffered seizures in infancy, 20% were described as irritable babies, and 10% as listless.

A global rating of motor development delay was calculated from the six items on the motor rating scale ($n=29$) (see Appendix K). A score of 0 to 6 was possible, ranging from no delay in motor development to delay in all
Table 10

<table>
<thead>
<tr>
<th>Percentage Indicating Problems on Parent Questionnaire</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Initial</td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>Child seems distant from the family</td>
</tr>
<tr>
<td>Child disrupts the family</td>
</tr>
<tr>
<td>What disciplinary approaches have been</td>
</tr>
<tr>
<td>most effective?</td>
</tr>
<tr>
<td>Spanking</td>
</tr>
<tr>
<td>Verbal discussion</td>
</tr>
<tr>
<td>Child's reaction to discipline:</td>
</tr>
<tr>
<td>Disregards</td>
</tr>
<tr>
<td>Rebels</td>
</tr>
</tbody>
</table>

six measured areas. The mean motor rating was 1.3 (SD = 1.5, range 0 to 5). Forty-one percent of the children were not delayed in any skill; 24% were delayed in one skill only.

Comparisons Between Groups

The subjects were divided into three groups at follow-up according to their diagnosis and CARS scores. Scores below 30 on the CARS are generally classified as nonautistic, scores between 30 and 36 are mild to moderate autism, and scores above 36, if five or more items are rated 3 or above, are classified as severe autism. CARS scores in
this sample had a mean of 31.4, and ranged from 19.5 to 45.0. One child with a score of 29 on the CARS was diagnosed autistic by both professionals making the concurrent diagnoses, and he was subsequently included in the autistic group. This resulted in 12 children in the nonautistic group, 15 in the mildly to moderately autistic group, and 8 in the severely autistic group. With one exception, all children diagnosed autistic by the CARS were diagnosed autistic by both persons making the concurrent diagnoses. The exception, noted above, was a child diagnosed autistic by the author and mentally retarded "with autistic features" by a psychologist.

Initial diagnoses at the clinic tended to be Developmental Delay or Developmental Language Disorder for non-autistic and mildly autistic children (see Table 11). Of 12 mildly/moderately autistic children, only 3 were originally diagnosed autistic. Two of those 3 were described as "untestable." Five of 7 severely autistic children were originally diagnosed autistic, and 4 of them were described as "untestable." Thus, severely autistic children and children who could not satisfactorily be tested were more likely to be called autistic at early ages (mean age for the clinic group at initial evaluation was 3.1 years).

One-way analysis of variance revealed no significant differences between groups on initial measures (see Table 12). The children were not significantly different on
Table 11

Initial Clinic Diagnoses by Group

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Non-autistic (n = 12)</th>
<th>Mildly/Mod. Autistic (n = 12)</th>
<th>Severely Autistic (n = 7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autism</td>
<td>0</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Developmental Delay</td>
<td>5</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Developmental Language Disorder</td>
<td>3</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Atypical Pervasive Developmental</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disorder</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cerebral Palsy</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mixed Specific Developmental Disorder</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>&quot;possible autism&quot;</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

the Autism Checklist and number of items checked on the Problem Checklist. Only one child in the severely autistic group had initially been given a PPVT-R, therefore that group was not included in the analysis of language scores. A t test revealed no significant difference between the nonautistic and mildly/moderately autistic groups on the language measure. In addition, there were no group differences on age of onset or age of initial evaluation. One-way analysis of variance revealed significant
Table 12

Mean Scores on Initial Measures by Group

<table>
<thead>
<tr>
<th></th>
<th>NVIQ</th>
<th>SQ</th>
<th>PPVT-R</th>
<th>Autism Checklist</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonautistic</td>
<td>72 (10)</td>
<td>75 (8)</td>
<td>65 (4)</td>
<td>7</td>
</tr>
<tr>
<td>Mild/mod autistic</td>
<td>52 (13)</td>
<td>60 (10)</td>
<td>57 (4)</td>
<td>9</td>
</tr>
<tr>
<td>Severely autistic</td>
<td>41 (4)</td>
<td>58 (4)</td>
<td>32 (1)</td>
<td>9</td>
</tr>
</tbody>
</table>

Note. Ns are shown in parentheses.

differences between groups on follow-up measures. Duncan's Multiple Range Test was used as a post hoc measure of pairwise differences between the three groups. See Table 13 for a summary of the group differences found.

Significant differences were found between the nonautistic and the two autistic groups on these measures taken at follow-up: nonverbal IQ ($F(2, 32) = 11.4$), Vineland SQ ($F(2, 31) = 23.97$), and the Autism Checklist ($F(2, 32) = 38.9$). The nonautistic group was significantly different from the severely autistic group on the Peabody Picture Vocabulary Test-Revised ($F(2, 24) = 8.19$). The Autism Checklist was the only measure that showed significant differences between the two autistic groups.

The distribution of follow-up IQ scores differs for the autistic and nonautistic groups (see Table 14). With 70% of
Table 13

Differences Among Groups on Follow-Up Measures

<table>
<thead>
<tr>
<th></th>
<th>NVIQ</th>
<th>SQ</th>
<th>PPVT-R</th>
<th>Autism Checklist</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonautistic</td>
<td>91.8</td>
<td>96.7</td>
<td>80.8</td>
<td>2.8</td>
</tr>
<tr>
<td>Mildly/mod. autistic</td>
<td>65.9</td>
<td>52.4</td>
<td>60.5</td>
<td><strong>6.9)</strong></td>
</tr>
<tr>
<td>Severely autistic</td>
<td>39.6</td>
<td>42</td>
<td>31.6</td>
<td><strong>9.1)</strong></td>
</tr>
</tbody>
</table>

a Significantly different from all other groups.
b Significantly different from severely autistic group.

As noted above, most children made gains on IQ tests, but there were group differences. In the nonautistic group, 10 of 12 children gained IQ points, with an average improvement of 14 points. Eleven of 15 mildly/moderately autistic children made IQ gains, averaging 12 points. Interpretation of the severely autistic group is hampered by the lack of initial IQ scores on half of the eight children. If these low numbers are averaged, the severely autistic group initially had a group IQ mean of 40.5 (n = 4) and at follow-up a mean of 39.6 (n = 8), suggesting no improvement.
Table 14

Distribution of Follow-Up IQ Scores by Group

<table>
<thead>
<tr>
<th></th>
<th>Autistic</th>
<th>Nonautistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal 90 and above</td>
<td>4 (17.4%)</td>
<td>5 (41.7%)</td>
</tr>
<tr>
<td>Low Average 80-89</td>
<td>3 (13.0%)</td>
<td>6 (50.0%)</td>
</tr>
<tr>
<td>Borderline 70-79</td>
<td>0</td>
<td>1 (8.3%)</td>
</tr>
<tr>
<td>Mildly Retarded 50-69</td>
<td>6 (26.1%)</td>
<td>0</td>
</tr>
<tr>
<td>Moderately Retarded 35-49</td>
<td>4 (17.4%)</td>
<td>0</td>
</tr>
<tr>
<td>Severely/Profoundly Retarded &lt;35</td>
<td>6 (26.1%)</td>
<td>0</td>
</tr>
</tbody>
</table>

Vineland SQs improved over time for the nonautistic group, but decreased for the autistic groups (see Table 15). In the nonautistic group the change was significant \( t(7) = -3.63 \). Their initial mean of 74.5, clearly in the delayed range of adaptive functioning, improved to an average score of 96.7. For the autistic groups, adaptive functioning scores decreased, but not significantly; the magnitude of the change was not as great as in the nonautistic group and it appears that a small \( N \) was a contributing factor in the failure to reach significance. When the two groups of autistic children are combined, Vineland SQs drop significantly \( t(13) = 3.5 \). Nine of 14 autistic children for whom initial Vineland scores were available decreased by 10 points or more at follow-up.
Table 15

Changes in SQs by Group

<table>
<thead>
<tr>
<th>Group</th>
<th>Initial Mean</th>
<th>Follow-Up Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonautistic (n = 8)</td>
<td>74.5</td>
<td>96.7</td>
</tr>
<tr>
<td>Mildly/Mod Autistic (n = 10)</td>
<td>60.1</td>
<td>52.4</td>
</tr>
<tr>
<td>Severely Autistic (n = 4)</td>
<td>58</td>
<td>42</td>
</tr>
</tbody>
</table>

*a Significant change over time.

It was reported that the number of autistic symptoms decreased over time; however, most of the change occurred in the nonautistic group (see Table 16). The decrease in symptoms in the mildly/moderately autistic group failed to reach significance, and there was no change in the number of symptoms for the severely autistic group.

Table 16

Change in Number of Autistic Symptoms

<table>
<thead>
<tr>
<th>Group</th>
<th>Initial Mean</th>
<th>Follow-Up Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonautistic</td>
<td>6.8</td>
<td>2.8</td>
</tr>
<tr>
<td>Mildly/Moderately Autistic</td>
<td>8.5</td>
<td>6.9</td>
</tr>
<tr>
<td>Severely Autistic</td>
<td>9.0</td>
<td>9.1</td>
</tr>
</tbody>
</table>

*a Significant change over time.
The chi-square statistical test was performed on each of the 14 items of the Autism Checklist. No group differences were found at initial testing, but some group differences emerged at follow-up. Nonautistic children were less likely to still be evidencing the following behaviors: (a) great difficulty in mixing with other children ($\chi^2(2) = 13.4$), (b) lack of fear about realistic dangers ($\chi^2(2) = 16.4$), (c) prefers to indicate needs by gestures ($\chi^2(2) = 9.5$), and (d) standoffish manner ($\chi^2(2) = 10.5$).

The chi-square test was used to examine group differences on the motor rating scale and the work/school functioning scale. No group differences were found for the motor rating scale ($\chi^2(8) = 5.63$). Most children showed little or no delay in their motor development; the ones who were delayed were equally distributed among the three groups. Group differences were found on the work/school scale. Table 17 shows the numbers from each group who attained particular ratings on the 5-point work/school scale. The five categories were collapsed to three in order to reduce the number of cells with expected frequencies below five. The two best outcome categories and the two worst outcome categories were combined. Outcome, as measured by the work/school scale, was significantly different for the three groups, $\chi^2(4, N = 35) = 22.32$. None of the nonautistic children were in the worst outcome category, and none of the autistic children were in the best
outcome category.

Table 17

Work/School Outcome Rating by Group

<table>
<thead>
<tr>
<th>Work/School Rating</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonautistic</td>
<td>2</td>
<td>3</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mildly Autistic</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Severely Autistic</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

Note. Outcome rating of 1 is best, 5 is worst.

The only measure taken at follow-up which did not discriminate between groups was the Problem Checklist. Children in all three groups showed equal numbers of these problems which are not distinct to autism.

Correlation Among Measures

Almost all variables were highly correlated with each other. Table 18 shows a correlation matrix of follow-up measures, all of which were significant. Note the strong negative correlation between CARS scores and the cognitive and adaptive measures; the more severe the autistic symptoms, the lower the IQ, PPVT-R, and Vineland SQ. Age of onset and motor delay were significantly correlated with follow-up NVIQ (r = .50 and -.46 respectively), but were not correlated with the other variables. The total score on the
Table 18
Correlations Among Follow-Up Measures

<table>
<thead>
<tr>
<th></th>
<th>SQ</th>
<th>PPVT-R</th>
<th>CARS</th>
<th>Work/School</th>
<th>Autism Checklist</th>
</tr>
</thead>
<tbody>
<tr>
<td>NVIQ</td>
<td>.74</td>
<td>.73</td>
<td>-.67</td>
<td>-.77</td>
<td>-.65</td>
</tr>
<tr>
<td>SQ</td>
<td>.80</td>
<td></td>
<td>-.77</td>
<td>-.86</td>
<td>-.76</td>
</tr>
<tr>
<td>PPVT-R</td>
<td></td>
<td>-.69</td>
<td></td>
<td>-.80</td>
<td>-.68</td>
</tr>
<tr>
<td>CARS</td>
<td></td>
<td></td>
<td></td>
<td>.74</td>
<td>.88</td>
</tr>
<tr>
<td>Work/school scale</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.74</td>
</tr>
</tbody>
</table>

Note. All correlations are significant.

Problem Checklist was the only follow-up measure that did not correlate with follow-up NVIQ.

Correlations were computed between initial and follow-up scores (see Table 19). Initial variables which were not correlated with outcome, as measured by the work/school scale, were: age at initial evaluation, age of onset of symptoms, motor delay rating, and Problem Checklist score.

Prediction of Outcome

The best initial predictor of follow-up status, as measured by the work/school scale, was the NVIQ. Multiple regression analysis using stepwise selection of variables failed to produce a combination of predictors superior to NVIQ alone. The addition of other variables (SQ, PPVT-R,
Table 19

Correlations Between Initial and Follow-Up Measures

<table>
<thead>
<tr>
<th>Initial Measures</th>
<th>NVIQ</th>
<th>PPVT-R</th>
<th>SQ</th>
<th>CARS</th>
<th>Work/ School Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>NVIQ</td>
<td>.79</td>
<td>.50</td>
<td>.74</td>
<td>-.59</td>
<td>-.68</td>
</tr>
<tr>
<td>PPVT (n = 9)</td>
<td>.74</td>
<td>.81</td>
<td>.75</td>
<td>-.60</td>
<td>-.71</td>
</tr>
<tr>
<td>SQ</td>
<td>.67</td>
<td>.54</td>
<td>.74</td>
<td>-.43</td>
<td>-.64</td>
</tr>
<tr>
<td>Autism Checklist</td>
<td>-.05</td>
<td>-.50</td>
<td>-.32</td>
<td>.28</td>
<td>.21</td>
</tr>
</tbody>
</table>

a Not significant.

Autism Checklist did not improve the ability of the equation to predict outcome.

The best predictors of autism severity at follow-up (CARS) were also examined. Among initial measures, the best predictor of autism severity was again NVIQ ($r = -.59$), and other variables (SQ, PPVT-R, Autism Checklist) did not enhance the regression equation.

Summary

Some important time-related changes and group differences were reported for a sample of autistic and autistic-like children. First, significant overall improvement in IQ and reduction in the number of autistic
symptoms were reported. Second, group differences, though not apparent at initial evaluation, emerged at follow-up. Nonautistic children made the most dramatic improvements in IQ. Vineland SQs improved for the nonautistic, but declined for the autistic children. Outcome was quite different for nonautistic as compared to autistic children. Lastly, the best predictor of outcome and of severity of symptoms at follow-up was the initial measure of nonverbal IQ.
CHAPTER IV

DISCUSSION

Thirty-five children were identified who had been diagnosed autistic or described as autistic-like. Initial test data and descriptions of behavior were gathered from records on the children, and a follow-up evaluation was done. Standardized tests for measuring nonverbal intellectual functioning, receptive vocabulary, and adaptive functioning were utilized, and the children were rated on measures of autistic behaviors. The CARS, which assesses the presence and severity of autism, was used to divide the children at follow-up into three groups: nonautistic, mildly/moderately autistic, and severely autistic. The resulting group of autistic children appeared quite representative of other diagnosed autistic children: 70% scored in the mentally retarded range (American Psychiatric Association, 1980), 57% were in the very poor outcome groups (Rutter, Greenfeld, & Lockyer, 1967), and the ratio of males to females was 10:1 (Achenbach, 1982).

When first evaluated at the clinic, all children were, on the average, 3 years of age. Intellectual, language, and adaptive functioning were in the retarded range, but all levels of functioning, from severely retarded to average, were represented. The children displayed an average of 8 of
the 14 symptoms on the Autism Checklist. Most were
developing normally in motor skills. They were difficult
for their parents to manage, and were likely to be described
as stubborn. Possibly the most significant finding of the
study was that none of these initial factors predicted
whether a child would be diagnosed autistic or nonautistic
at follow-up. At initial evaluation, children who would
later be classified nonautistic did not score significantly
higher on language or nonlanguage tasks, were not
functioning higher in adaptive skills, were not displaying
fewer or different sets of autistic symptoms, and were not
developing faster motorically. Even "untestability" was as
likely to occur in the nonautistic as in the mildly/
moderately autistic (two cases each).

Follow-up evaluations occurred an average of four years
later. There were important changes revealed in the test
scores, and most of the changes were among the nonautistic
group. The following discussion will highlight the most
important time-related changes and group differences which
were found. This will be followed by a discussion of the
results as support for the view of autism as a spectrum
disorder. Finally, the findings of certain infant
characteristics which were revealed in the study will be
discussed in relation to research in key areas of child
development.
Time-Related Changes

Before emphasizing group differences, it is important not to lose sight of the overall changes that occurred with time for most of the children. The number of autistic symptoms decreased for all but the severely autistic. Nonverbal intellectual functioning improved dramatically for over half of the children, increasing enough to place them in a higher IQ category. Social difficulties were still widespread at follow-up, with 69% displaying a "standoffish manner" and 66% "great difficulty in mixing with other children," but the children as a whole were not the severe discipline problems that they had been at younger ages.

The most striking changes occurred in the nonautistic group. At follow-up, their intellectual and adaptive functioning were in the low average range. The nonautistic group showed, on the average, only 3 of the 14 symptoms on the Autism Checklist. They were the only group represented in the two highest categories on the work/school outcome measure: only nonautistic children were functioning in regular classrooms or with minimal special education assistance (e.g., speech therapy). Furthermore, none were represented in the two worst outcome categories which require intensive special education services and close supervision in all activities.

In contrast to the nonautistic group, 70% of the autistic children were functioning in the mentally retarded
range, even on essentially nonverbal tests. Sixty-four percent showed a decrease in their Social Quotient by 10 points or more. More than half were in the two worst outcome categories on the work/school scale. Among severely autistic children, there appeared to be no improvement in scores: IQ, adaptive functioning, language scores and autistic symptomatology all remained stable or declined over time.

The following will provide a more detailed discussion of each of the areas where changes occurred: autistic behaviors, cognitive functioning, adaptive functioning, and behavior management. Implications for clinical practice and limitations of the study will be highlighted where appropriate.

**Autistic Behaviors**

The Autism Checklist was the only initial measure of autistic symptoms available on the children. It is generally used as a screening measure, and a score of 7 is often taken as a cutoff score. The number of autistic symptoms shown by the children decreased over time for all but the severely autistic, a finding consistent with the developmental literature. Overall, mean scores on the Autism Checklist decreased significantly from 8 to 5.9; the decrease was most dramatic for the nonautistic group whose means fell from 7 to 3.

Even on initial evaluation, 13 of the children did
not score a minimum of 7. The initial Autism Checklist was scored retrospectively from the medical and psychological reports written on each child when first evaluated. If any of the initial reports mentioned a behavior described on the Autism Checklist, that behavior was checked as present. Consequently, the method used in rating subjects on the Checklist may not be reliable for assessing the number of autistic behaviors that were actually present at the time of the initial evaluation. It is possible that some behaviors were present but not specifically mentioned in the reports. The fact that many subjects failed to reach an arbitrary score of 7 may therefore be due to the method used in scoring the Checklist.

The follow-up Autism Checklist, on the other hand, was scored based on observation of and interaction with the child, in addition to an interview with the parent regarding present behaviors. Given that the follow-up Autism Checklist was probably a more accurate assessment of autistic symptoms than the initial Autism Checklist, the decrease in the number of symptoms becomes even more impressive.

The CARS reflected more autistic symptomatology at follow-up than the Autism Checklist. The mean CARS score of 31.4 was in the autistic range. The CARS appears to be less sensitive to developmental changes (i.e., a child retains his or her autistic status despite predictable changes due
to maturation). It has been shown to be a more reliable instrument than the Checklist for identifying autism in an older population of children and adolescents (McCallon, 1986).

The CARS would almost certainly have been a better assessment of the presence of autistic symptoms at initial evaluation. It is possible that an initial CARS score would have significantly discriminated differences in the children, and would have predicted eventual diagnosis and outcome whereas the initial Autism Checklist did not. Five clinic children had initially been evaluated using the CARS, but since the number was very small and the reliability with the present examiner unknown, these scores were not utilized.

Checklists are easy to use, even for people who have not been specifically trained to administer them, for example teachers or other school personnel. They are appropriate as screening devices to identify children who need a more thorough evaluation, which is how the Autism Checklist is primarily used. When a more thorough examination is needed a measure like the CARS becomes necessary. It is harder to administer than a checklist, and requires professional training to do so, but it provides much more comprehensive assessment of the child's functioning. A checklist does not, and is not intended to, predict diagnosis (the initial Autism Checklist in this
study was not significantly correlated with follow-up measures). The results of this study suggest that when the Autism Checklist is used as a screening instrument, a lower cutoff score should be considered, especially if the child is school-age.

From the results of this study and the developmental literature, it appears that most young children who can be described as autistic or autistic-like will show a diminishing number of those symptoms as they reach the school years. Caution should be taken in counseling parents about this trend, however, because it is not known to what extent early intervention contributed to these improvements. The time-related changes found do not necessarily mean that these changes occur automatically with development. The children in this study were identified early, and thus were more likely to receive intervention, especially in the form of early childhood educational programs. In the absence of intervention, we do not know what course the symptoms might have taken, thus predictions of diminishing symptoms must be contingent upon early intervention.

Cognitive Functioning

The trend for scores on nonverbal intelligence tests to increase over time was dramatic. While the unreliability of early IQ scores, particularly from infant scales, is well-documented (Sattler, 1982), the findings with autistic children have not been so definitive. IQ scores of autistic
children have been described as being remarkably stable (Achenbach, 1982). In Rutter's follow-up study (1969; 1970) IQ was found to change, but the change was as likely to be down as up. DeMyer et al. (1973) found nonsignificant IQ changes that tended to be upward for children with higher IQs and tended to decrease, or remain the same, for children who were globally retarded. Two important factors may serve to reconcile these findings with the present results. First of all, the average age at initial evaluation in both Rutter's and DeMyer's studies was more than 5 years. By the age of 6 or 7, IQ scores become more reliable (Sattler, 1983). Second, in both studies initial IQ was inadequately assessed and often the Vineland SQ was used as a measure of IQ. Results from the present study suggest that the SQ in autistic children can decline appreciably over time. Such a decline may have tended to counterbalance any increase in scores on more acceptable measures of intelligence.

A more recent study by Freeman, Ritvo, Needleman, and Yokota (1985) reported stability in IQ scores for young autistic subjects. The authors tested a group of autistic children yearly on standardized intelligence and language tests. Initial ages ranged from 2 to 6 years with a mean of 3 years, 10 months. Performance IQ scores were reported for 48 children initially. Preliminary findings suggest that cognitive and linguistic test scores of young autistic children remained stable over time. The criterion they used
to assess stability was movement between two categories: retarded (IQ scores below 70) and nonretarded (IQ scores of 70 and above). In their study, only eight children (13%) changed IQ groups, and all who did were under 3 years of age at initial testing.

Applying this criteria for change to the present study, 4 of 23 autistic children (17%), changed from retarded to nonretarded functioning on intelligence tests. This figure compares more closely with the 13% in the Freeman et al. study.

Thus, depending on the definition of change, IQ scores of young autistic children may be described as stable, but for purposes of clinical application, Freeman et al.'s two-category criteria appears too limited. Using their system, for example, it would be possible for a child with an IQ of 49 and a child with an IQ of 70 to improve by 20 points and not change groups; their scores would be described as stable. A great deal of individual improvement is subsequently ignored. By the same token, a child with an IQ of 68 who improved by 2 points would change groups, giving a clinically insignificant change apparent significance.

For the purposes of describing the reliability of IQ scores in young autistic children, it seems more informative to use more than two descriptive categories. The differences between severely, moderately, and mildly retarded functioning are not insignificant, and the
distinctions influence our predictions of outcome. To refer to the "stability" of cognitive scores in young autistic children is a generalization, not applicable to individual cases.

The clinical implications of changing IQ scores cannot be overemphasized. Children with severe developmental problems such as the ones in this study are most often seen and tested very early in life, and even though, on the average, lower scores tend to be more reliable over time than higher ones, it is not a hard and fast rule that can be applied to individual cases seen in a clinic. This sample showed some IQ gains that were no less than remarkable. Ten children showed gains of 20 or more points, five children 30 or more, and one child (mildly/moderately autistic) showed a gain of 53 IQ points. Among the eight "untestable" children, all were testable at follow-up; two scored above 100, one scored 79, and two scored in the 60's. With IQ so essential to these children's prognoses, it is imperative to keep in mind, and to communicate to the parents, the instability of early scores.

The language measure employed in this study was the Peabody Picture Vocabulary Test--Revised, a test of receptive vocabulary that does not involve expressive language. Only nine initial test scores were available, so it is difficult to assess change. Based on these small numbers, all three groups of children had initial PPVT-R
scores in the retarded range of functioning. By the time of follow-up, the nonautistic group had improved to a low average mean of 81 while the two autistic groups continued to function in the moderately to severely retarded range. Also at follow-up, all of the nonautistic group had sufficient language ability to take the test, but eight of the autistic children did not.

Since the test does not involve expressive language, more autistic children will be successful on it than on a test of language production. The drawback is that the PPVT-R is not a comprehensive language assessment. It tells us nothing about a child's use of language. One of the crucial aspects of language impairment in autism is the failure to use language for communication (Rutter, 1979). Studies of autistic children using measures of expressive language generally find language scores significantly lower than nonverbal IQ scores (cf. DeMyer et al., 1973), whereas in this study, nonverbal IQ and receptive language scores were similar. The important findings here are the same as for IQ; that is, initially, nonautistic and mildly/moderately autistic children scored similarly low on a measure of receptive vocabulary, and it was not until follow-up that the groups became distinguishable, with the nonautistic group improving and the autistic groups showing no improvement.
Adaptive Functioning

Vineland SQs improved significantly for the nonautistic group, and decreased significantly for the two autistic groups combined. These opposite trends are probably explainable by the nature of the Vineland Social Maturity Scale itself. What is being measured by the Vineland items changes over the age periods of the test (Sattler, 1982). Children without the handicap of autism will progress from the lower levels of self-help behavior measured in the infant and early childhood items to the higher levels of self-directed activities and communication measured by later items. Autistic children can and do learn simple self-help skills (e.g., getting a drink unassisted), and certainly most have no problem negotiating the independent motor skills measured in the early part of the test (e.g., walks about the room unattended). As they grow older, their handicap increasingly affects their ability to pass Vineland items. They may never "participate in pre-adolescent play" (social age = 8.28) or "go about home town freely" (social age = 9.43). Thus, in comparison to chronological age, their social age is not developing at a consistent rate, and their social quotient actually drops over time. This is what occurred with the autistic children in this study. More than half decreased their SQ scores by ten points or more, while the nonautistic group made striking improvements that paralleled their improvements in intellectual and
language development.

None of the major follow-up studies reviewed above reported SQ scores from both initial and follow-up evaluations, so it is not known if other groups of autistic children have shown similar trends on the Vineland. Regardless, it would be interesting to look at the revised Vineland Adaptive Behavior Scales (Sparrow, Balla, & Cicchetti, 1984) longitudinally with this population of children. The original Vineland showed some large gaps in item difficulty as the ages increased (Gordon & Reed, 1985), and the SQ has the same disadvantages as a ratio IQ (Sattler, 1983). Either of these factors could have affected the autistic children's scores. The Survey Form of the revised edition has two and one-half times the number of items as the original scale and it produces standard scores; it may therefore have overcome these disadvantages. The revised edition also produces more than one scale score, as opposed to the single SQ. It may be that the deficits of autistic children will be more clearly defined, for example, their scores in communication and socialization might decline, for example, but they might be keeping pace in the areas of daily living skills or motor skills. Whether autistic children would still show a decline in adaptive functioning scores as they grow older, awaits investigation. A cursory look at items from the revised edition suggests that the same trend may be found. The Daily Living Skills
domain still contains an expectation for a growing financial independence, awareness of danger and ability to communicate (e.g., via telephone), that are not characteristic of autistic children. Autistic children just do not keep up the consistent pace toward independence that is the criterion for adaptive functioning.

**Behavior Management**

A Parent Questionnaire assessed the behavior management difficulties encountered by parents of autistic and autistic-like children. Problems such as disruption to the family, rebellion against discipline, and consistency of disciplinary techniques were addressed by the questions. The percentages reported for these items are descriptive of the present sample of children only. As the forms are unique to the clinic, no baseline data are available on a nonclinic population; therefore, a conservative approach should be taken, drawing only those conclusions which seem most obvious.

It is apparent from the questionnaire that at younger ages, these children were a significant behavior management problem for their parents. Most parents used spanking as a disciplinary approach, but approximately half of the children either disregarded parental discipline or rebelled against it. Furthermore, despite the severe language impairment of their children, many parents tried verbal discussion as a disciplinary measure as well. It appears
that the parents attempted to discipline their child as parents of nonhandicapped children might, but with this special group of children, the methods were ineffective. Given the high frequency with which parents checked "stubborn" as a description of their child, it is possible that parental expectations are high compared to the cognitive abilities of the children.

Over time these children became less of a management problem for their parents. Only a small percent of the children continued to disrupt the family, disregard discipline, or rebel. The parents, for their part, used less spanking and more verbal discussion. This trend is quite developmentally appropriate for children in general, but, again, becomes questionable given the language handicap of this population. The parents, in large numbers, continued to describe their children as "stubborn."

These findings point to the need for behavior management training for parents, especially during the preschool years. Severely handicapped children are a major stress to a family, and autistic children present even more difficulties than other handicapped children (Holroyd & McArthur, 1976). Parents of autistic children need management skills which do not depend on a high, or even average, level of language ability in the child. They need help in formulating reasonable expectations for their child. Re-framing the child's response to discipline may be
helpful. A child who does not understand a verbal command is not being stubborn when he or she fails to comply.

To summarize the discussion of changes which occurred over time, autistic behaviors decreased for all but the severely autistic and nonverbal intellectual functioning improved for most of the children. Receptive vocabulary and adaptive functioning scores showed different trends depending on group membership. All children were difficult management problems for their parents, although that too improved with time.

The changes which consistently occurred with these children between initial and follow-up evaluation highlight the importance of frequent re-evaluation of this population. Whether it is to assess changes in intellectual functioning, development in language skills, changes in adaptive functioning, or to evaluate the appropriateness of diagnosis or school placement, frequent re-evaluation in the preschool years is essential. Educational planning and program development depend upon accurate assessment of ability. School placement is often determined by diagnosis. These are not small matters, rather they are crucial to the optimum development of the child. Special education classes, along with behavior therapy, are probably our most powerful treatments for autistic children (Rutter, 1985). Parents are sometimes reluctant to have their child frequently tested and school districts may have arbitrary
re-evaluation periods which are too lengthy (e.g. every three years). With young children, the possibility of changing status and the importance of monitoring that status need to be clearly explained to parents so that they can make informed decisions about re-testing.

From the standpoint of clinical implications, perhaps the single most important finding of the study is the similarity that nonautistic and mildly/moderately autistic children showed on initial evaluation. Several measures were examined and none provided discrimination of the two groups. This suggests that when a 3-year-old is evaluated and found to display symptoms of mild/moderate autism, it is not possible to predict what that child will be like at five years of age. Intellectual functioning in the retarded range can improve to average or remain retarded. Adaptive functioning in the retarded range can do the same, and can even decline with respect to age. Symptoms of autism will probably decrease, but it is uncertain whether they will decrease enough so that a diagnosis of autism is inappropriate. As satisfying as it may be to tell parents we know what is wrong with their child and we know what can be expected of him or her in the future, we must instead tell them that we cannot be sure. Early diagnoses in these cases appear to be as unreliable as early IQ scores. Fortunately our recommendation is likely to be the same regardless of whether the child is eventually diagnosed
autistic or not. What either the autistic or autistic-like child at age 3 needs is intensive special education efforts directed toward the language and social deficits they share. Each should benefit from such intervention.

It is not possible to say that the changes and improvements which occurred over time were the result of educational intervention or merely maturation. The children in the study were, on the whole, identified early, and thus more likely to receive early intervention. While many of the children became involved in early childhood educational programs in the area, no investigation was made into the quantity or quality of those programs. The developmental literature on autistic children has reported diminishing symptomatology as the child grows, but the findings of impressive IQ improvement and changing adaptive functioning scores await replication. The children in this study may not be representative of the general population in terms of the improvements seen, especially in the absence of early intervention.

In regard to the generalizability of results, two further cautions are in order. First, the children in the study were diagnosed on the basis of DSM-III criteria, which has been praised for its specificity (Achenbach, 1982) and reliability (Schopler, 1983). Since the study was undertaken, however, the American Psychiatric Association has revised the criteria for the diagnosis of autism in
favor of a less specific system that does not require age of onset prior to 30 months of age (DSM-III-R). Clinicians and researchers will be using a different set of criteria in the future, and it is not yet known what impact this will have on the identification of autistic children.

Second, in combining the clinic and school samples of autistic children, some group homogeneity was lost. While the differences between the groups were not statistically significant, in part due to small sample sizes, the school group was consistently lower on initial measures of nonverbal IQ and adaptive functioning, with respect both to the mean and the range of scores (refer back to Table 5). The advantages of combining the groups must be considered in light of the homogeneity that was lost.

The Severity Continuum

Autism is generally viewed as occurring along a continuum of severity, and this study lends support to that idea. The differences that mildly/moderately autistic children showed as compared to the severely autistic appeared to be quantitative rather than qualitative (i.e., the milder group scored somewhat higher on cognitive and adaptive measures, but showed the same trends of declining scores on the Vineland and inability to function in the less restrictive classrooms representative of a better outcome).

The question of whether the autistic-like children fall somewhere along the same continuum cannot be unequivocally
answered by the present findings. Just as the severely affected child has the same underlying syndrome as the mildly affected one, so they may also have the same syndrome as the child who falls just below the cutoffs on various diagnostic criteria. When evaluated as young children, they do not show impressive differences, and it can be argued that the differences they show later are due almost entirely to differences in intellectual ability. The same factors, severity of the disorder and IQ, greatly influence the prognosis of autistic children. Much effort has gone into a search for the boundaries of autism and for the essential defining criteria, when, in fact, the syndrome may be characterized more by overlap than by definitive boundaries. Shea and Mesibov (1985) have hypothesized a relationship between severe learning disabilities and higher-level autism. Certainly many of the nonautistic children in this study would qualify for a diagnosis of learning disability based on discrepancies between nonverbal and verbal skills.

Many of the children in the nonautistic group may have the same syndrome, in much milder form, as the autistic children. Eight of twelve nonautistic children in this study were diagnosed either Developmental Language Disorder or Atypical Pervasive Developmental Disorder. In a study designed to look at the questionable end of the continuum, these are arguably the most questionable in terms of representing anything but a very mild end of the autism
spectrum. The other four nonautistic children appear to belong in other distinct syndromes. For example, the one with Schizoid Personality Disorder appears to be developing symptoms of schizophrenia; this child has already been hospitalized once for psychotic symptoms which included hallucinations. Such symptomatology separates him from autistic children. Separation Anxiety Disorder, also represented among the nonautistic group, does not necessarily suggest a mild form of autism. Cohen, Paul, and Volkmar (1987) propose that several syndromes are represented by atypical children, some of which are related to autism and some of which are not.

The findings of striking differences between autistic and autistic-like children at follow-up, however, seem to argue strongly that the two groups are different. The nonautistic children made dramatic improvements in every area while overall gains eluded the children in both autistic groups. At follow-up, mildly/moderately and severely autistic children looked much more like each other than like the nonautistic group.

Even if there was strong evidence that most autistic-like children did have a very mild version of autism, it is not proposed here that they should be called autistic. Their prognosis appears to be so much better that to label them autistic might only serve to erect a barrier to their progress. The term autism carries with it a heavy load of
dire predictions about future social ineptness and dependent functioning. For purposes of early intervention, however, they have similar educational needs: small, structured classes, language development, and social skills training, for example.

Findings of Infant Characteristics

Over half of all the children in the study were described as "not cuddly as a baby," and a smaller number were described as irritable or listless as infants. While not a primary focus of this study, such findings would appear to have implications for two important areas of research in child development, namely temperament (Thomas & Chess, 1977) and attachment (Ainsworth, Blehar, Waters, & Wall, 1978).

Temperament is usually studied longitudinally, starting with infants who have no known psychological disorder. For example, the best known study of temperament is one done by the New York Longitudinal Study which has followed 133 subjects from early infancy to early adulthood (Thomas & Chess, 1986). The authors have identified three temperamental constellations. The easy child is characterized by regularity, positive approach to new stimuli, adaptability to change, and generally positive mood. There are two categories of difficult children. One is characterized by irregular biological functions (e.g. feeding), negative withdrawal responses to new stimuli,
difficulty in adapting to changes, and intense mood expressions that are frequently negative. The other type of difficult child is called the slow-to-warm-up child; initial negative responses to new stimuli gradually give way to adaptability after repeated contact. Either of the two difficult categories sound as if they could be describing an autistic child; indeed it is hard to imagine an autistic child falling into the easy category at all. Thomas and Chess have found that children with the difficult child pattern are more vulnerable to the development of behavior problems in early and middle childhood, and that this risk increases if the child is mentally retarded or physically handicapped. Their findings seem to have implications for autistic children, but given the nature of the longitudinal design and the incidence of 4 autistic children per 10,000 in the population, temperament studies have not directly addressed children with this disorder. Temperament could, however, impact on attachment, and attachment responses have been studied in autistic children.

Studies of attachment behavior generally employ a set of staged episodes described by Ainsworth et al. (1978) which observe a child's reactions to being separated from his or her mother, being with a stranger, and then being reunited with the mother. It is usually necessary to vary the procedures somewhat with samples of handicapped youngsters (Blacher & Meyers, 1983). In a group of 14
hospitalized autistic children, Sigman and Ungerer (1984) observed attachment behaviors in the form of proximity-seeking during reunion with the mother and a differential amount of social behavior directed toward the mother as opposed to the stranger. The autistic children did not show distress during separation from the mother, although the fact that they were hospitalized and already separated from the mother surely had some impact on response to separation. This study had a small number of children, and half of them did not demonstrate observable attachment behaviors.

Shapiro, Sherman, Calamari, and Koch (1987) investigated attachment behaviors in a group of developmentally disordered children that included 15 with autism, 10 with Atypical Pervasive Developmental Disorder (APDD), 8 with Developmental Language Disorder, and 3 with mental retardation. They found secure attachments in half of the autistic and half of the APDD children. Quality of attachment did not correlate with diagnosis. Furthermore, the attachment behaviors shown by these developmentally disordered children were not significantly different from the attachment behaviors shown by normal children at a younger age.

It has been speculated that autistic children are unable to form attachments (cf. Blacher & Meyers, 1983), but again empirical research may be able to dislodge a myth about autism. Autistic children "are not uniformly aloof
and isolated" (Shapiro et al., 1987, p. 483). It would be helpful in the future to identify how attachment is accomplished in spite of the severe handicaps present in the disorder. Perhaps from the standpoint of the parent we could learn how they facilitate attachment with babies who resist being held and who avoid eye contact. There may be a repertoire of skills in such parents that can be taught to others who are finding it difficult to respond to their infant. Attachment in infancy sets important groundwork for later peer relationships (Sroufe, 1983). Autistic and autistic-like children in this study continued to show major deficits in social development; for the autistic child, we know that the deficits will last a lifetime. Because of the importance of attachment, again, early intervention with parents seems called for.

Summary

Autistic-like children were included in a follow-up study of autistic children in order to examine the similarities and differences between the two groups. Autistic-like children have generally been neglected by research which has strived to clearly define its autistic populations by weeding out the difficult-to-diagnose children.

First, on initial evaluation, at an average age of 3 years, nonautistic children were very similar to mildly/moderately autistic children on measures of nonverbal
intellectual functioning, receptive vocabulary, adaptive functioning, and numbers and kinds of autistic symptoms.

Second, important changes occurred over time, and the most dramatic changes were in the nonautistic group. At follow-up, nonautistic children were functioning in the average or near-average range on tests of nonverbal IQ, adaptive functioning, and receptive vocabulary. The number of autistic behaviors that they still evidenced fell to an average low of three. In contrast, autistic children, while most improved their IQ scores, showed declines in adaptive functioning scores and no improvement on the language test. Severely autistic children showed the most stability of any group, apparently making no improvement on IQ, language, or adaptive functioning scores, or in numbers of autistic behaviors displayed.

These changes and group differences were discussed in terms of a developmental trend for autistic symptoms to diminish and in relation to other reports of the stability of IQ scores. It was emphasized that even among autistic children, IQ scores at early ages can and do change. Implications for clinical work were highlighted. The management challenge these handicapped children present for their parents, especially at early ages, was discussed and recommendations were made for early intervention and parent training. The changes in adaptive functioning were discussed in terms of the nature of the Vineland Social
Maturity Scale. All of the changes found in the study highlight the need for frequent re-evaluation in this population.

The nature of the syndrome of autism was discussed, particularly in its relation to the milder, atypical children who were included in this study. Support was found for the concept of autism as a spectrum disorder whose milder forms are related to and overlap with the more severe forms; however, the findings of superior follow-up status for the autistic-like children raise serious questions about including them in the same syndrome as the autistic children.

Finally, the difficult nature of the children as infants was discussed in relation to research on temperament and attachment. A suggestion was made for future work in the area to illuminate how parents of difficult babies facilitate attachment, in the hope that identifiable skills could be taught to others, and an impact could be made on the pervasive social deficits.
APPENDIX A

Nine-Point Criteria of the British Working Party
APPENDIX A

Nine-Point Criteria of the British Working Party

1. gross and sustained impairment of emotional relationships with people;
2. apparent unawareness of his own personal identity to a degree inappropriate to his age;
3. pathological preoccupation with particular objects or certain characteristics of them, without regard to their accepted functions;
4. sustained resistance to change in the environment and a striving to maintain or restore sameness;
5. abnormal perceptual experience (in the absence of discernible organic abnormality);
6. acute, excessive and seemingly illogical anxiety as a frequent phenomenon;
7. speech either lost, or never acquired, or showing failure to develop beyond a level appropriate to an earlier age;
8. distortion in motility patterns;
9. a background of serious retardation in which islets of normal, near normal or exceptional intellectual function or skill may appear (Creak, 1963, p. 84).
APPENDIX B

DSM-III Criteria for Infantile Autism
APPENDIX B

DSM-III Criteria for Infantile Autism

A. Onset before 30 months of age.

B. Pervasive lack of responsiveness to other people (autism).

C. Gross deficits in language development.

D. If speech is present, peculiar speech patterns such as immediate and delayed echolalia, metaphorical language, pro-nominal reversal.

E. Bizarre responses to various aspects of the environment (e.g., resistance to change, peculiar interest in or attachments to animate or inanimate objects).

F. Absence of delusions, hallucinations, loosening of associations, and incoherence as in Schizophrenia.
APPENDIX C

Childhood Autism Rating Scale
APPENDIX C

Childhood Autism Rating Scale

Abnormality

None    Mild    Moderate    Severe
1        1.5     2         2.5     3         3.5      4

1. Relationships with people
2. Imitation
3. Emotional responses
4. Use of body
5. Relation to objects
6. Adaptation to environmental change
7. Visual attention
8. Listening attention
9. Touch, taste, and smell
10. Fear or nervousness
11. Talking
12. Pointing and gesturing
13. Activity level
14. Inconsistencies in intelligence
15. General impression
APPENDIX D

Autism Checklist
APPENDIX D

Autism Checklist

1. Great difficulty in mixing and playing with other children.
2. Acts as deaf. No reaction to speech or noise.
3. Strong resistance to any learning, either new behavior or new skills.
4. Lack of fear about realistic dangers—may play with fire, climb dangerous heights, run into busy road or into the sea.
5. Resists change in routine. Change in the smallest thing may result in acute, excessive or seemingly illogical anxiety—child rejects new or all but a few foods.
6. Prefers to indicate needs. Speech may or may not be present.
7. Laughing and giggling for no apparent reason.
8. Not cuddly as a baby. Either holds himself still or clings limply.
9. Marked physical overactivity. Child may wake and play for hours in the night and yet be full of energy the next day.
10. No eye contact. Persistent tendency to look past or turn away from people especially when spoken to.
11. Unusual attachment to a particular object or objects. Easily preoccupied with details or special features of this object, and has no regard for its real use.

12. Spins objects, especially round ones. Can become totally absorbed in this activity and distressed if interrupted.

13. Repetitive and sustained odd play—flicking pieces of string, rattling stones in a tin, tearing paper.

14. Standoffish manner. Communicates very little with other people. Treats them as objects rather than people.
APPENDIX E

Problem Checklist
APPENDIX E

Problem Checklist

1. Is always on the go or acts as if "driven by a motor."

2. Has difficulty staying seated in school, at church, meals, etc.

3. Difficult to get his/her attention.

4. Has difficulty keeping attention (concentration) on school work or other tasks until completed.

5. Often does not seem to listen.


7. Needs a lot of supervision (one on one attention).

8. Has difficulty in groups, play, or school.

9. Frequently disrupts class.

10. Often speaks or acts before thinking.

11. Has difficulty waiting turn in games or group activities.

12. Sleep problems (resists going to sleep, wakes up in night, restless).

13. Explosive temper or tantrums.


15. Stubborn.

16. Provocative behavior.

17. Difficult to discipline.
18. Destructive.
19. Difficulties in getting along with other children.
20. Socially withdrawn (prefers to be alone).
22. Excessive fears or worries.
23. Seems unhappy or depressed.
24. Doesn’t like self.
25. Overly dependent on parents, siblings.
26. Reports unusual or peculiar thoughts.
27. Runaway behavior.
28. Sexual misbehavior.
29. Lies excessively.
30. Stealing.
31. Involvements with drugs, alcohol, etc.
32. May injure other persons or self.
33. Delinquent (involvement with police)
34. Seems to have characteristics of opposite sex.
35. Others.
APPENDIX F

Parent Questionnaire
APPENDIX F

Parent Questionnaire

Underline correct answers or fill in correct blank.

1. Child (disrupts, gets along with, seems distant from) family.

2. Discipline has been (strict, lenient, inconsistent, effective, consistent).

3. What disciplinary approaches have been most effective? (spanking, withholding privileges, grounding, verbal discussion)

4. Child's reaction to discipline: (disregards, rebels, accepts and follows through, effective for a short time).
APPENDIX G

Work/School Rating
APPENDIX G

Work/School Rating

Class I -- Normal

Preschool: Learn counting, drawing, colors, etc. as well as agemates. If not in pre-school, parents' description of activities strongly indicate normal intellectual skills.

School-age: Learn in school all subjects—reading, writing, arithmetic—as well as or better than an average child.

High school: Learn all subjects as well or better than average teenager. Earn some spending money.

Post-high school: In college or technical training or holding a job without difficulties. Self-supporting or nearly so if holding a full-time job.

Class II -- Borderline Normal

Preschool: Have some difficulty in keeping up with learning of agemates but can be managed in regular class. If not in preschool, parents report slight delay in learning counting, drawing, colors.

School-age: In class with normal children but behind age level in one or more subjects or one or two years behind grade level. If retained in one grade, passed second time.

High School: Difficulty with academically demanding subjects. Completing high with difficulty or may have
dropped out. To be given this rating, person must have been in normal high school at least a few months.

Post-high school: Functioning satisfactorily in job that makes few intellectual or organizational demands. Low salary. No college. Housekeeping marginally satisfactory.

Class III -- Educable Retarded

Preschool: Can't learn as much as agemates in nursery school class; may be in special education early childhood class; children have some splinter intellectual skills that approximate age level at school or at home.

School-age: In class for retarded or autistic children where child learns some intellectual subjects at level lower than average. If started in normal school at age six, was either removed as not ready to start or retained but failed to learn and ultimately had to be removed from normal school classes.

High school: In special classes, but has learned to read and also some writing.

Post-high school: Must have much training and special assistance to function in non-demanding job. Doing well in sheltered work shop. Capable of performing some household chores without supervision. Cannot maintain own home or apartment without supervision but can go places in the community alone.

Class IV -- Trainable Retarded

Preschool: In class for trainable retarded or autistic
children; have no sign of budding intellectual skills but possess at least one perceptual-motor skill at age level demonstrated in home or school.

School-age: Can function in class for trainable retarded.

High school: In special classes, cannot read or write. Requires some supervision in going about the community alone.

Post-high school: Can work only in sheltered work shop but requires much assistance to do this. Can do simple tasks at home under supervision.

Class V -- Subtrainable Retarded

Preschool: Cannot learn in preschool. Have no perceptual-motor skill even near age level.

School-age: Cannot function in class for trainable retarded.

High school age and beyond: Cannot work or do any household tasks, even with supervision. Must be taken by another person for excursions into community.
APPENDIX H

Letter Announcing Research Project
APPENDIX H

Letter Announcing Research Project

The Child Study Center is doing follow-up evaluations of some of the children we have seen during the last few years. This is a joint research project we are conducting with North Texas State University. We hope it will help us better understand children with developmental problems. We are interested in knowing more about how these children develop after the initial evaluations.

The evaluation will consist of interviewing you and administering some tests with your child. The entire evaluation will take about 2 hours. There will be no cost to you. A brief summary of the results of the evaluation will be available to you to help in planning for your child. Denise McCallon, a graduate student at North Texas, will be calling you to arrange an appointment. We hope you will participate in this research project as we feel it will benefit all of us who are involved, including you and your child.

Thank you very much for your time and consideration. If you have any questions, please feel free to call me at 336-8611.

Sincerely,

Harry Brick, Ph.D.
Director of Psychological Services

Note: This letter was on Ft. Worth Child Study Center letterhead.
APPENDIX I

Letter From Special Education Department

Recommending Participation in Research
Dear Parents:

The Fort Worth Child Study Center is doing follow-up evaluations of children with autism. It is a joint research project they are conducting with North Texas State University. This study is one that I would support.

The evaluation consists of interviewing the parent and administering some tests with your child at the Fort Worth Child Study Center, 1300 W. Lancaster. The evaluation takes about 2 hours and there is no cost to you. A summary of your child's test results will be sent to you.

If you are interested in participating, please call the Fort Worth Child Study Center, Psychological Services, at 336-8611. Leave your name and telephone number with the secretary and you will be contacted to set up an appointment.

I feel that the results of this study could be helpful in educational planning for the students involved. Thank you for your time and consideration.

Sincerely,

Kay Jones
Director of Special Education
APPENDIX J

Informed Consent
APPENDIX J

Informed Consent

Child's Name

I hereby give consent to Denise McCallon to give psychological tests to my child and use the results of the testing in her research project. The testing and interviewing procedure will take approximately two hours. I have heard a clear explanation and understand the nature of the tests to be used. I understand that the benefit to me and my child will be an increased understanding of my child's current functioning.

After the study has been completed and the results analyzed, and if I desire, I may receive information regarding those results. If I have any questions regarding the procedures or results I may contact Denise McCallon at the Child Study Center (336-8711) or her university supervisor, Deborah Garfin, Ph.D. (817/565-2673).

I understand that I may withdraw my consent at any time. Having received this information and satisfactory answers to my questions, I voluntarily consent to the testing and the use of the results for research.

SIGNED: ____________________________ SIGNED: ____________________________
WITNESS PERSON RESPONSIBLE

Date ____________________________  RELATIONSHIP
APPENDIX K

Ages (months) at Which the Motor Items of the Bayley Scales of Infant Development Are Passed
APPENDIX K

Ages (months) at Which the Motor Items of the Bayley Scales of Infant Development Are Passed

<table>
<thead>
<tr>
<th>Item</th>
<th>Percentiles</th>
</tr>
</thead>
<tbody>
<tr>
<td>13. Sits with support</td>
<td>1.0 2.3 5.0</td>
</tr>
<tr>
<td>28. Rolls from back to stomach</td>
<td>4.0 6.4 10.0</td>
</tr>
<tr>
<td>29. Sits alone, steadily</td>
<td>5.0 6.6 9.0</td>
</tr>
<tr>
<td>38. Stands up by furniture</td>
<td>6.0 8.6 12.0</td>
</tr>
<tr>
<td>45. Stands alone</td>
<td>9.0 11.0 16.0</td>
</tr>
<tr>
<td>46. Walks alone</td>
<td>9.0 11.7 17.0</td>
</tr>
</tbody>
</table>
REFERENCES


Bartak, L., & Rutter, M. (1976). Differences between mentally retarded and normally intelligent autistic


Autism Rating Scale: A validity and reliability study.
Paper presented at the meeting of the Southwestern Psychological Association, Ft. Worth, TX.


Sattler, J. M. (1982). *Assessment of children's intelli-
gence and special abilities* (2nd ed.). Boston: Allyn and Bacon.


