This is to certify that the

dissertation entitled

BEHAVIORAL AND COGNITIVE ANTECEDENTS
OF READING ACHIEVEMENT IN THE ELEMENTARY GRADES:
A LONGITUDINAL STUDY

presented by

ROBERT L. JARVIS

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in EDUCATION

Harvey F. Clarizio
Major professor

Date JULY 13, 1982

BEHAVIORAL AND COGNITIVE ANTECEDENTS OF READING ACHIEVEMENT IN THE ELEMENTARY GRADES

By

Robert L. Jarvis

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Counseling, Educational Psychology,
and Special Education

1982

ABSTRACT

BEHAVIORAL AND COGNITIVE ANTECEDENTS
OF READING ACHIEVEMENT IN THE ELEMENTARY GRADES

By

Robert L. Jarvis

The purpose of the study was two-fold: 1) to identify teachers' perceptions of behavioral attributes of elementary age students which were highly and significantly related to reading achievement over time, and 2) to investigate the efficiency and utility of using teachers' ratings of student behavioral attributes to predict future reading performance. A descriptive, longitudinal design was instigated whereby behavioral ratings were evaluated independently as well as in conjunction with IQ, initial level of reading achievement, and sex of student, as to their ability to predict subsequent reading achievement after one, two, and three year intervals.

Behavioral attributes found to be most highly related to reading achievement over time appeared to reflect a student's degree of adaptation to the task orientation of the classroom, degree of self-motivation and persistence, and degree of general social maturity. Appropriate behavior adjustment was found to be a generally more important factor in terms of reading success for boys relative to their female peers. All of the behavior scales developed in the study were found to be significantly related to reading achievement over time. Results further indicated that despite the strong linear relationship between the behavior scales and future reading achievement, teachers' ratings were not found to be as predictive of later achievement as were scores on a measure of general

intelligence and a measure of initial level of reading achievement.

Generated prediction equations which utilized only the behavior scales in estimating future achievement tended to greatly underpredict the performance of above-average readers and greatly overpredict the achievement of below-average readers. Further, despite significantly high multiple correlations, as well as high diagnostic "hit rates," the generated prediction equations which incorporated initial level of achievement, IQ, and sex of student were found to be generally lacking in terms of accuracy in predicting above-average and below-average readers over time.

The findings of the study are finally presented in light of their relationship to past research, their implications for implementation of reading screening and diagnostic programs, and their implication for future research.

With much love and pride I dedicate this effort to Nancy. Her on-going expression of love, patience and support continues to encourage me to reach for heights I never though possible.

ACKNOWLEDGEMENTS

Thanks to members of my committee, Dr. Harvey Clarizio, Dr. Bill Frey, Dr. Gaston Blom, and Dr. Richard Prawat, who have provided on-going support and encouragement throughout my graduate school experience. Special thanks to Dr. Frey, whose guidance and support in both academic and non-academic domains greatly expanded my experience and horizons while at MSU.

Further thanks to all of my dedicated friends and relatives who labored through the preparation of various drafts of this document -- Kac, Karen, Sue, Mary, Barbara and Roberta.

I am forever indebted to Bob and Mary Ann who graciously provided the gift of friendship as they opened their hearts and home to me over the past years during completion of this project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

THE PROBLEM

Introduction

Significant interest and concern on the part of both educators and parents regarding the reading achievement of elementary school-age children necessitates a critical examination of potentially relevant factors which may contribute to students' relative success or failure in reading.

Measures of cognitive functioning, as well as sex and socio-economic status have been shown to share a significant amount of variance with the reading achievement scores of elementary age children. However, psychologists and teachers alike have concurrently been well-sensitized to the fact that "non-intellectual" factors, e.g. emotional state, personality, and overt classroom behavior, should, and most likely do, play a vital role in a child's academic life.

The relationship between social-emotional functioning and reading achievement is at present quite unclear despite a significant body of research spanning the past fifty years. While there is an implicit assumption within much of the research that social-emotional functioning is not the primary causal factor in the behavior-achievement relationship (Allyon and Roberts, 1974; Gates, 1968; Spache, 1976), the direction

of the relationship remains empirically unresolved. It is further difficult to interpret consistent trends in the relationship from past empirical efforts due to the relatively diverse conceptions and subsequent measurements of "non-intellectual" functioning and academic achievement. Many studies have been further limited by narrowness in the child population sampled, especially in terms of potentially relevant variables such as IQ, sex, and age or grade level.

Particularly noticeable is the relative absence of longitudinal-developmental perspective with regard to the relationship between emotional/behavioral factors and reading achievement in the normal population of elementary age students. Few studies have investigated the relative long-term efficiency of using "non-intellectual" factors in predicting reading achievement. The majority of studies have focused on concurrent comparisons of student characteristics and achievement. Results of the few longitudinal studies reported appear somewhat inconclusive with regard to the relative contribution of cognitive and "non-intellectual" factors to reading attainment. However, all would suggest that emotional or behavioral factors do make a significant and unique contribution to our understanding of the nature of childrens' reading achievement in the elementary grades (Feshback et al., 1977; Kohn and Rosman, 1974; Lambert and Urbanski, 1980).

Improving our understanding of the predictive relationship between student classroom behavior and reading achievement has potential value from a number of viewpoints. At a general level, one could expect that teachers, counselors, and psychologists may gain from an increased understanding of the dynamics of classroom behavior and its relative contribution to reading success or failure during the early course of reading development; a point argued by Peck (1971) who called for an increased

sensitivity and awareness of educators to the coping effectiveness of pupils in the classroom. Along the same line, Camp and Zimet (1974) argue the need for identification of positive, achievement-enhancing behavior, and the need for clarification of whether a lack of reading achievement is related more to the absence of positive classroom behaviors or the presence of negative, disruptive behaviors. The literature is replete with studies of the negative behavioral and personality characteristics of disabled readers. Behavioral variables have been shown in many studies to make a significant, unique contribution to the variance in reading achievement in the elementary grades. Such findings would appear to be important in light of the early identification and subsequent remediation of learning and/or emotional problems (Bryan and Wheeler, 1972; Feshbach, et al., 1977; Lambert and Nicole, 1977). Kellam, et al. (1975) for example, found that even when intelligence test scores and school achievement were included in a prediction battery, ratings of children's social adaptation status in first grade made a significant unique contribution independent of intelligence to the prediction of later need for remedial help and grade retentions in the intervening years.

The work of Cobb and his associates suggests the practical relevance of investigating the relationships between classroom behaviors and academic achievement in terms of direct treatment of underachievement. In an initial series of studies (Cobb, 1970; Cobb, 1972a) specific behavioral categories, which appeared to be significant predictors of achievement, were defined. In samples of first and fourth grade children, Cobb found a consistent relationship between performance on standardized measures of achievement and such behavioral classes as attending to teacher, compliance with teacher requests, and volunteering to answer

academic questions. Subsequent studies (Cobb and Hops, 1973; Hops and Cobb, 1974) involved efforts to develop these "survival skills" in underachieving pupils through the use of behavior modification procedures. They found that these intervention procedures were effective in increasing and maintaining survival skill behaviors and that such gains were directly related to gains in reading achievement.

Need for the Study

Investigators in psychology and education have worked for many years to develop methods of predicting students' levels of reading achievement. As has been indicated, despite a significant body of research spanning the past fifty years, the predictive relationship between emotional-behavioral functioning and reading achievement is far from being clarified. During this time, most of the research has focused upon deliniating specific personality traits of disabled readers. These attempts to identify a consistent personality profile of the disabled reader have proven essentially futile. The early conclusions of Gates in his classic 1941 review of the relationship appear to have withstood the test of time:

1. Personality difficulties are frequently but not universally associated with reading difficulties.
2. In cases where they occur together, personality difficulties may be causes, concomitants, or results of reading difficulties. The direction is not clear.
3. There is no single personality pattern characteristic of reading failure and there is no proved one-to-one relationship between type of adjustment difficulty and type of reading disability. For example, feeling of insecurity resulting from

> undue home pressure for achievement may result in low reading achievement marked by withdrawal; compulsive, anxious reading marked by frequent errors; or it may result in higher achievement in reading than would be expected from mental level.

Bower and Holmes (1959), nearly two decades later, reviewed over one hundred subsequent studies of the personality-reading achievement relationship, and indicated, ". . . in the main, the same general conclusions arrived at by Gates are, but for slight modifications still applicable." They found that only a few investigations reported significant differences for specific personality traits between good and poor readers, and that many such differences tended to be "nicely counterbalanced" by other studies which reported differences in exactly the opposite direction.

Review of these earlier studies indicated that the clarification of the relationship between personality and achievement was significantly limited by a general lack of success in defining and measuring personality (Russell, 1953); Spache, 1976). This validity issue is critical, as the ability to establish consistent relations between "non-intellectual" factors and reading achievement depends to a significant extent upon the meaningfulness of the personality or behavioral categories used.

Others have argued that discovery of such relationships is essentially irrelevant for use by the regular classroom teachers for improving a child's reading level (Bateman, 1966; Engleman, 1967). Over the past fifteen years, however, there has been an increased emphasis upon the study of more discrete, observable classroom behaviors, which may be potentially under the teacher's control, and their relationship to reading achievement.

It appears that two major areas of research have contributed to this increased emphasis on the study of overt classroom behaviors rather than personality traits, which might be predictive of academic achievement. The first, as alluded to previously, was the reported inconsistency in linking specific personality traits with reading success or failure (Bower and Holmes, 1959; Cobb, 1969; Gates, 1941; and Russell, 1953). For example, Cobb (1969) reviewed ninety-one doctoral dissertations reported in Dissertation Abstracts between 1961 and 1969 which examined the relationship between academic achievement and such variables as self-concept, anxiety and other personality traits. He found, as had earlier reviewers, that while modest correlations were discovered in several of the studies, replications of the findings were often difficult to obtain.

Concurrently, in the late 1960's, a number of researchers were investigating means of altering inappropriate classroom behavior through procedures based upon social learning theory principles. The majority demonstrated marked improvements in the classroom behavior of emotionally disturbed children e.g. (Hewett et al., 1967; Hops, 1971; Patterson et al., 1969) and disruptive children (Hall, et al., 1968; Madsen et al., 1968). These studies indicated that children's inappropriate classroom behavior could be significantly altered, however the question remained whether changes in behavior were concommitant with positive changes in academic achievement. Indeed, there was no clear evidence in any of the aforementioned studies that achievement had improved.

Similarly, efforts to improve academic achievement by treating emotional disorders through traditional means, e.g. psychotherapy and counseling, appeared not to have been exceptionally fruitful (Ashcraft, 1970; Cobb, 1972; Munger, et al., 1964). During this same period of time, Lipe and Jung (1971) suggested a need for a clear and empirically

demonstrated rationale relating classroom behavior to measured achievement.

Promising findings began to emerge as researchers began focusing on the relationship of observable student behaviors, as opposed to personality dimensions, with reading and mathematics achievement in elementary school pupils. In a five year follow-up study, Meyers, et al. (1968) demonstrated significant relationships between behavior ratings in kindergarten and achievement scores in the fifth grade. "Attention", emerged as the first and most powerful predictor, even among other verbal ability measures, of all aspects of fifth grade reading achievement, including reading words, reading comprehension, and spelling. The correlation of "attention" to each of these subtests was .40, .43, and .38 respectively. Using direct observation procedures in sixth grade classrooms, Lahaderne (1968) similarly found significant correlations between rate of attending behavior and reading achievement on a Scott-Foresman Basic Reading Test and the Stanford Achievement Test. Correlations were .51 and .46 respectively for boys and .49 and .39 for girls.

A review of relevant studies over the past fifteen years which have subsequently investigated more specific non-intellectual/behavioral correlates of reading achievement in elementary school children indicates rather consistent findings regardless of often diverse methodologies and instrumentation for measuring behavior and reading achievement. Three clear behavioral categories have emerged as being significantly related to early reading achievement: <u>independent</u>, <u>task-oriented behavior</u> (Coker and Lorentz, 1977; Engin, 1975; Kohn and Rosman, 1974; McKinney, et al., 1975; Soli and Devine, 1976); <u>attending behavior</u> (Cobb, 1970, 1972; Feshbach, et al., 1977; Forness, et al., 1977; Meyers, et al., 1968; Samuels and Turnure, 1974; Soli and Devine, 1976; Swift and Spivack, 1967);

and compliance with teacher and classroom demands (Camp and Zimet, 1974; Cobb, 1970, 1972; Coker and Lorentz, 1977; Lambert and Nicoll, 1977; Soli and Devine, 1976).

These findings are significant in that they suggest that not all undesirable student classroom behaviors are necessarily related to poor reading achievement. Behaviors which teachers may consider most undesirable because they are disruptive of classroom order, e.g. verbal or physical aggression, may in fact, be of least importance to early reading success (Blom, et al., 1980; Cobb, 1972; Lambert and Nicoll, 1977; Lambert and Urbanski, 1980). Such findings further lend credence to future studies which attempt to differentiate between academically-related behavioral attributes or dimensions and those behaviors which may more accurately reflect general social adjustment or mental health issues. As Lambert and her colleagues have suggested, one of the major limitations of past research is that it has not generally focused on the "relevant" student characteristics related to reading success.

As has been suggested, the prediction of reading achievement and the early identification of children who are at high risk for school failure have been of concern to educators for many years. The majority of efforts have focused upon the development of predictive indices or instruments composed of cognitive and student behavioral variables, both independent from and in conjunction with one another. These instruments have for the most part been developed for a preschool or kindergarten population with the intention of screening for eventual reading failure.

Stevenson, et al., (1976) suggested, however, that an understanding of the relative utility of these various measures have been generally difficult to attain. This situation would appear to be primarily due to the fact that the majority of predictive studies cited have tended to

rely solely on multiple correlation as the primary statistical procedure. Although this type of analysis indicates the relative degree of association between the predictors and the criterion, it provides limited information as to the instrument's practical value. In particular, it does not indicate how many or what types of errors in prediction are made. Practical use of such screening instruments may convey unnecessary cost to schools as well as risk to an individual child if the utility of the device is inadequately assessed (Satz and Fletcher, 1979).

In practice, psychologists and educators frequently make important decisions concerning the classification, prognosis, and treatment of individual students, wherein treatment decisions are made on the basis of early test signs. The accuracy of these predictions, however, has been rarely addressed. As Satz and his colleagues suggest, a true error rate of prediction is frequently masked by obvious confoundings from misclassification of high risk subjects (false negatives) or low risk subjects (false positives), making accurate judgments regarding student outcomes tenuous. The false negatives refer to subjects for whom good achievement was predicted and whose actual achievement was "not good," while false positives refer to those subjects for whom poor achievement was predicted and whose actual achievement was "not poor." The confoundings of misclassification appear to be well exemplified in the studies of Stevenson, et al. (1976) wherein relatively high and significant multiple correlations were found between various cognitive tasks in pre-kindergarten and subsequent reading achievement in the first through third grades (.66 to .77), however the ability of the predictive index to discriminate among good and poor readers in the third grade was relatively deficient. The results of this and other predictive studies strongly suggest the need for future research to go beyond a purely concurrent correlation model, towards a longitudinal

perspective which assesses both the accuracy as well as practical utility of a predictive scale.

Satz and Fletcher (1979) suggest that beyond the notion of accuracy the predictive utility of a measure implies that detection signs (e.g. valid positives and negatives) should be evaluated within the broader context of the prevalence estimates of the event to be predicted - in this situation the base rates of student reading success or failure. Meehl and Rosen (1967) concur and further suggest that the efficiency of the great majority of psychometric devices are essentially impossible to evaluate because:

1) Base rates are virtually never reported
2) Most devices are reported without cross-validation data
3) Results are frequently reported only in terms of significance tests for differences between groups rather than in terms of the number of correct decisions for individuals within the groups.

These shortcomings of past studies must be addressed in future investigations in order to gain a better understanding of the predictive relationship between "non-intellectual" student characteristics and subsequent achievement in reading.

Purpose of the Study

The purpose of this study was essentially twofold: 1) to identify teachers' perceptions of behavioral attributes of elementary age students which were highly and significantly related to reading achievement over time, and 2) to investigate the efficiency and utility of using teachers' ratings of student behavioral attributes to predict future reading performance. The study incorporated a descriptive, longitudinal design, whereby teachers' ratings of student behavior were evaluated independently and concurrently with IQ, initial level of reading achievement, and sex

of student, as to their ability to predict subsequent reading achievement after one, two, and three year intervals.

The design of the present study is developed in light of what have been identified to be limitations of past studies which have proported to investigate the predictive relationship between student characteristics and academic achievement. As has been suggested, past studies have generally based their findings upon a single sample in which predictor and criterion variables were assessed at a single point in time, and relative amount of variance to reading achievement attributed to the cognitive and behavioral factors reported. Due to a lack of longitudinal perspective in these studies, little empirical evidence is available as to the relative importance of student classroom behavior to future reading success or failure - an important point of knowledge in terms of a psychologist's or educator's decision to target or not target an individual student for special support services, e.g. special education, remedial reading.

Closely related to the aforementioned problem is the need to determine the accuracy by which future reading achievement may be assessed by behavioral and/or cognitive predictors. Few studies have reported the practical value of their instruments in terms of actual predictive ability and efficiency. This study assesses the diagnostic utility of a predictive instrument composed of behavioral characteristics related to reading achievement relative to base rates of reading ability and disability.

Overview of the Study

Chapter I has presented the background, rationale, purpose, and objectives which serve as the basis for the investigation. In Chapter II the pertinent related research is reviewed. In Chapter III the research method is described, including sample characteristics, instrumentation, data collection procedures, statistical hypotheses, data analysis

procedures, and definition of key terms. Chapter IV includes the presentation of data and evaluation of the testable hypotheses. Results are discussed in Chapter V in light of past research and practical implications, considering limitations and delimitations of the study. Chapter VI offers a summary of the study and recommendations for future research.

CHAPTER II

REVIEW OF THE LITERATURE

This chapter has been divided into four main sections which reflect the major issues and concerns in past studies which have investigated the prediction of reading achievement in elementary aged students. The first section focuses upon the reported relationship between student social-emotional/behavioral functioning and subsequent reading achievement. The research reported primarily reflects those studies which have shown a high concurrent correlation between these variables. As student behavior has been evaluated primarily by two means, teacher ratings and classroom observation, this review is similarly dicotomized.

In the second section, longitudinal, predictive studies of reading achievement are reviewed. This section addresses the major issues regarding the accuracy and practical utility of prediction of reading achievement. As sex of student has been found to be significantly related to both classroom achievement and behavior, the third section of this review investigates the reported differences between elementary aged boys and girls in both reading achievement and classroom behavior. Finally, the fourth section examines the evidence of the validity of teachers' judgments regarding elementary student behavior.

Student Classroom Behavior and Reading Achievement

A review of studies over the past fifteen years which have subsequently investigated behavioral-affective correlates of reading achievement in elementary school children indicated rather consistent findings regardless of often diverse methodologies and instrumentation for measuring behavior and reading achievement. Three clear behavioral categories have emerged as being significantly related to early reading achievement: independent, task-oriented behavior (Coker and Lorentz, 1977; Engin, 1975; Kohn and Rosman, 1974; McKinney, et al., 1975; Soli and Devine, 1976; attending behavior (Cobb, 1970, 1972; Feshbach, et al., 1977; Forness, et al., 1977; Meyers, et al., 1968; Samuels and Turnure, 1974; Soli and Devine, 1976; Swift and Spivack, 1967); and compliance with teacher and classroom demands (Camp and Zimet, 1974; Cobb, 1970, 1972; Coker and Lorentz, 1977; Lambert and Nicoll, 1977; Soli and Devine, 1976).

Teacher Ratings of Student Behavior and Reading Achievement

Swift and Spivack (1967) investigated the relationship between teacher ratings of pupil behavior and students' academic success or failure in grades kindergarten through six. The purpose of their project was to learn how disturbing classroom behaviors were organized in both normal and special education classrooms, and to assess the relationship between these behaviors from the Devereux Elementary School Rating Scale and report card grades. The ranges of simple correlations for each of the ten factors of the Devereux scale and teacher marks in reading across the elementary grades were:

1)	Classroom Disturbance	-.21 to -.37
2)	Impatience	-.10 to -.54
3)	Disrespect/Defiance	-.08 to -.27
4)	External Blame	-.13 to -.27
5)	Achievement Anxiety	.00 to -.27
6)	External Reliance	-.27 to -.77

7)	Inattentive/Withdrawn	-.15 to -.74
8)	Irrelevant Responses	-.07 to -.40
9)	Creative Initiative	.21 to .60
10)	Need Closeness to Teacher	-.02 to .50

Analysis of individual factors by grade indicated that in the first years of school the child who approaches and responds to the teacher in a warm and friendly manner is more likely to be academically successful. Ability to attend and work independently similarly were indicative of successful readers. Regarding differences between the sexes, boys were rated as presenting significantly more problems than girls on nine of the ten factors. Girls were rated as being more inclined to approach and be friendly toward the teacher, a significant predictor of early academic success. Further, as a group, boys were rated as being more disturbing and disruptive, more impatient, and more disrespectful. The authors concluded that boys clearly demonstrated more "achievement-impeding behaviors," and that girls are significantly more often judged as conforming to or capable of meeting classroom behavioral demands.

Ratings on two teacher rating scales were compared with observations of classroom behavior and reading achievement by Camp and Zimet on the Pittsburgh Adjustment Survey Scale (PASS) and the Connors Teacher Rating Scale (CTRS). Reading achievement was derived from the reading subtests of the California Achievement Test and Metropolitan Achievement Test as well as the Gates-MacGinitie Reading Test, and Gilmore Oral Reading Test. Results indicated that classroom observation of "off-task" behavior correlated significantly and positively (r=.31) with the Daydreaming subscale of the CTRS and significantly but negatively (r=-.38) with the Prosocial subscale of the PASS. The measures of reading achievement were found to be significantly positively correlated with the Prosocial subscale of the PASS (r=.56 to .81), negatively with the Daydreaming subscale (r= -.59 to -.76), and negatively with the total CTRS (r= -.30 to -.58).

Engin (1975) examined the relative importance of academic aptitude, as measured by the Wechsler Intelligence Scale for Children, and behavioral variables as measured by teacher ratings on the Devereux Elementary School Behavior Rating Scale on classroom achievement of third and fourth grade inner-city pupils. Analysis of the regression of the reading-related subtests of the Stanford Achievement Test on the WISC and Devereux Scale indicated that both intelligence and behavioral variables constituted important, independent components of the reading achievement of inner-city third and fourth graders. Behavioral factors contributing significantly to the explained variance included external blame (r=-.26), creative initiative (r=.30), and impatience (r=-.17), a profile similar to that found by Swift and Spivak (1967). One significant anomoly in Elgin's findings was the positive contribution (r=.39) of disrespect/defiance to the prediction of the Word Meaning subscale of the Stanford. She postulated that for "inner-city" children this dimension may be reflective of an active verbal orientation to school may have "an enhancing rather than deleterious effect" upon aspects of reading achievement.

Kohn and Rossman (1974) found in their sample of over two hundred lower and middle class boys in New York City that preschool teachers ratings of social-emotional behavior were significantly predictive of second grade achievement. A Task-Orientation dimension of the Schaefer Classroom Behavior Inventory correlated .33 and .44 with the Word Knowledge and Reading subtests of the Metropolitan Achievement Test. This dimension tapped the child's ability to concentrate, become absorbed, and persevere in classroom activities.

Harrison (1976) evaluated two hundred underachievers on the Wechsler Intelligence Scale for Children (WISC) and the Wide Range Achievement Test (WRAT), and obtained teacher ratings of pupil behavior on a Classroom Behavior Checklist. "Distractibility" was found to be the behavior most

consistently associated with low reading and spelling achievement as measured by the WRAT. The interaction between distractibility and reading and spelling achievement was most significant with the younger, early elementary students.

Twenty-seven disabled readers were rated by teachers on the twenty categories of the Burks Behavior Rating Scales (BBRS) in an investigation by Wilf (1972). The subjects were further divided into two subgroups on the basis of the severity of reading retardation. Results indicated that the severity of reading disability appeared to be significantly related to certain characteristics of maladjustment, that is "poor ego strength" and "poor attention" appeared to be much more characteristic of the severely disabled group.

Feshback, et al. (1977) conducted a five year longitudinal study which focused on the comparison of psychometric versus behavioral kindergarten predictors of reading performance in the first through third grades. During the spring of their kindergarten year over 800 children in middle class Los Angeles Schools were administered a variety of individual tests including the Wechsler Preschool and Primary Scale of Intelligence (WPPSI), the de Hirsch-Jansky Predictive Index, and Bender Visual-Motor Gestalt Test. During the same period pupils were rated by their teachers on the Student Rating Scale (SRS). The primary criterion variables consisted of standardized achievement tests, the Cooperative Primary Reading Tests and a criterion-referenced individualized reading inventory, both administered in the first through third grades.

In the first grade the total SRS score proved to be the most predictive of the cognitive and behavioral independent variables. For second grade reading achievement, SRS remained the best predictor with an average correlation of .50, closely followed by the de Hirsch and IQ

with correlations of .47 and .40 respectively. Behavioral ratings remained consistently high in third grade. Even with IQ partialed out, the correlation of SRS with third grade achievement was .33. Multiple regression analysis indicated that SRS and IQ accounted for 51% of the total variance in third grade achievement.

Sex of the child appeared to be a significant moderating variable in the prediction of achievement by use of both behavioral and cognitive variables. Several interesting differences in the pattern of correlations were found between boys and girls. IQ was consistenly a better predictor of both reading and mathematics achievement for boys than for girls. In contrast, while the SRS and the de Hirsch were more highly correlated with first grade reading achievement in boys than in girls, by the third grade the correlations for girls were higher than those found in boys. This pattern held for all factors of the SRS including an attention-behavioral control factor, but excluding a language dimension which operated very much like IQ. The authors suggested that motivational variables may be more important for girls in the early primary grades than IQ, and at the same time, kindergarten behavioral factors, which reflected poorer performance by boys than girls, may decline in importance for boys with increasing maturity.

Behavioral dimensions derived from the Pupil Behavior Rating Scale is a study of Lambert and Nicoll (1977) proved to be potentially useful predictors of first and second grade reading achievement. They found that an "adaptation to classroom demands" factor was a strong predictor of both first and second grade reading achievement ($r=.60$, $r=.66$), as measured by the Cooperative Primary Test. This factor included such attributes as following directions, working independently, and interest in learning tasks. These attributes, individually, were also significantly

related to reading achievement: i.e. difficulty following directions (r=-.53, r=-.63), dependent on teacher (r=-.51, r=-.52), no interest or enthusiasm for learning (r=-.57), 5=-.63). Distractibility was also found to be highly correlated with first and second grade reading achievement (r=-.51, r=-.50).

Lambert and Nicoll further found that the adaptation dimension combined with the demographic variables of sex, race and socioeconomic status was more predictive of reading achievement than was cognitive functioning in combination with the demographic indicators. These results were consistent with the earlier findings of Lambert (1972) and Lambert, et al. (1976) which found that these same behavioral attributes measured in the second and fifth grades were as good, and in some cases better, predictors of subsequent ninth and twelfth grade achievement as were intelligence and achievement test data from the elementary grades.

Final analysis of Lambert and Nicoll's (1977) study indicated (a) the powerful relationship of the classroom adaptation dimension to early reading; (b) the weaker, but significant relationship of an intrapersonal factor, i.e. unhappy, withdrawn, behavior with early reading; (c) the relative unimportance of acting-out, aggressive interpersonal behavior with reading achievement. The latter finding is interesting in light of the similar findings of Elgin (1975) with inner-city pupils and Blom, et al. (1980) with lower elementary aged girls.

Lambert's findings regarding the strong relationship between "adaptation" and achievement were further confirmed in a recent study in which second, fourth, and sixth grade subjects were assigned to twelve behavior typology groups based upon teacher's ratings from the Pupil Behavior Rating Scale (Lambert and Urbanski, 1980). A one way analysis

of variance procedure comparing reading scores across the twelve groups was highly significant and post hoc comparisons again suggested the strong association between those attributes of the classroom adaptation dimension and reading achievement. Differences in measures of the interpersonal and intrapersonal dimensions did not appreciably contribute to the observed group differences in achievement.

Observed Classroom Behavior and Reading Achievement

Studies over the past fifteen years, which have investigated the relationship between observed classroom behavior of students and achievement, have generally found three dimensions of behavior which appear to contribute a substantial amount of variance to reading achievement: on- and off-task behavior, attending behavior, and compliance with teacher demands.

In a major replication of the early correlational studies of Meyers et al., (1968) and Lahaderne (1968) which found a relatively high relationship between "attention" and reading achievement, Cobb (1970) demonstrated that greater relationships could be obtained if the behavior categories of the predictors were more precisely defined. He suggested that the earlier aforementioned studies utilized ratings which were too global and which included a number of irrelevant behavior's that masked the true relationship between specific, discrete behavior responses and achievement.

Cobb began a series of studies in 1969 to clarify the relationship between a combination of specific student behaviors and achievement (Cobb, 1969, 1972a). In the first study, over one hundred fourth graders were observed in five classrooms from two schools during arithmetic instruction over a ten day period. Fourteen discrete behaviors such as

attending to task, talking to teacher about academic material, volunteering, compliance with teacher requests, and out of seat behavior, etc. were recorded. The rate of each child's behavior was then used as an independent variable in producing, through a step-wise procedure, multiple regression equations to predict attainment as measured by the Stanford Achievement Test. The results indicated that the specific behavior designated as attending, talk to peers about academic material, self-stimulation, out-of-chair, and compliance were significantly related to both mathematics and reading achievement, with out-of-chair and talk to peers about academic material being the two best predictors of reading achievement.

In a subsequent study, Cobb (1970) used similar procedures in eight first grade classrooms where students were observed in both reading and arithmetic activities. The major predictors of reading achievement were found to be attending (r=.59) and volunteering (r=.59), and for arithmetic, looking around and compliance. Further, attending and looking around, which were as might be expected highly correlated with each other (r=-.75), were found to be significant predictors of both reading and mathematics attainment.

When Cobb (1970) split his sample into male and female subgroups, it was found that observed classroom behavior was more highly predictive for boys than girls. The multiple correlation was .71 for boys versus .45 for girls. One major difference in terms of the specific predictors, was the finding that for boys compliance and noncompliance behaviors were significant predictors of achievement, whereas for girls no significant relationship for these behaviors was found.

Samuels and Turnure (1974) further investigated sex differences in classroom attention and its relationship to reading achievement at the

first grade level. A behavioral observation schedule was used in which 74 boys and 58 girls were observed during reading instruction. Task-relevant behaviors, e.g. looking at text or teacher were considered inattentive. Interrater reliability was reported at .89. Reading achievement was measured by recognition of forty-five words randomly selected from the Dolce List of Basic Sight Words and total attending scores for each pupil were placed into four quartiles.

Results of the study indicated no significant differences on a reading readiness measure between boys and girls, and these readiness scores were used a covariate in later analysis. It was found that mean word recognition scores were significantly higher for girls (30.03) than for boys (22.68), and that these scores tended to increase from the lower to the upper quartiles of attention. Further, overall mean attention scores significantly favored the girls, and significant sex differences were found for each of the four attention quartiles all favoring the girls. An overall correlation of .44 between attention and reading achievement was found which was consistent with the finding of Cobb (1972) for fourth graders (r=.45) and by Lahaderne (1968) for sixth graders (r=.51 to .39).

Forness, et al. (1977) focused on observable kindergarten behavior to determine its utility in predicting eventual educational risk. Subjects were forty children followed through the end of second grade. At that time, school records were examined and information on special status such as retention or referral to remedial reading or special education were noted. On the basis of this information, all subjects were divided into two groups, risk and nonrisk. Correlation coefficients were then computed between the kindergarten observation categories of positive peer

interaction, task attention, nonattention, and disruptive behavior and eventual second grade status. Significant differences were found between the two groups with attending behavior being the highest predictor (r=.47) with eventual risk.

McKinney, et al. (1975) investigated the predictive value of a number of observed behavioral patterns in regard to fall and spring academic achievement for ninety second graders. Students were observed for five minute periods on each of four days during language arts instruction in the fall of the school year and again in the spring. Twenty-seven discrete classroom behaviors were systematically coded at ten second intervals for each of the students. Interrater reliability was reported at .90. Factor analysis procedures identified twelve behavioral categories which were retained as independent variables for predicting achievement on the California Achievement Test.

Using multiple regression procedures, it was determined that the following four categories of fall behavior predicted 33 percent of the variance in fall achievement: distractible behavior, passive responding, dependency, and constructive play. High frequencies in the first three were associated with lower initial achievement and high frequencies of constructive play (appropriate self-directed activity, not part of the curriculum) were predictive of higher initial achievement. In the spring frequencies of constructive self-directed activity (independent classwork) and aggression were found to be significant predictors of spring achievement. The final multiple R for predicting spring achievement from fall hehavior was .60.

Analysis of sex effects failed to show significant differences in fall behavior patterns for boys and girls, however, in the spring girls

evidenced a higher frequency of constructive self-directed activity than did boys.

The general portrait of the academically competent child attained in this study appears to be similar to the previously reported findings (Cobb, 1972; Lahaderne, 1968; Samuels and Turnure, 1974). The results of McKinney et al. indicated that the child who is attentive, independent and task-oriented in his classroom interaction with his peers is more likely to succeed academically than the child who is distractible, dependent on teacher for direction, and passive in peer group activities.

Soli and Devine (1976) utilized the observational procedures developed by Cobb (1970, 1972a) in their investigation of the relationship between observed behavior across academic subjects and reading achievement as measured by the Gates-MacGinitie Reading Test. Twenty-five observations were made on each of 312 third and fourth graders. Multiple regression procedures indicated that total achievement among all students was best predicted by the behaviors positive peer interaction, not attending, self-stimulation, and initiation of activity with teacher, giving a final regression coefficient of .45. When the total group was divided into two achievement levels (high and low achievers), it was found that in the high group achievement was best predicted by positive peer interaction, initiation of activity with the teacher, inappropriate locale, and noisy with a final regression coefficient of .44. Achievement in the low group was also predicted with a final coefficient of .44, however, the behavioral predictors were different from those of the high group; not attending, self-stimulation, playing, complying.

Coker and Lorentz (1977) examined the moderating influence of different classroom settings, Teacher-Directed versus Program Directed,

upon the relationship between observed pupil classroom behavior and reading achievement. Six hundred students in grades three through eight were observed during both the fall and spring of the academic year by means of the Coping Analysis Schedule for Educational Settings (CASES), and reading achievement was assessed by means of the comprehensive subtest of the Iowa Tests of Basic Skills.

Stepwise regression procedures indicated that both fall and spring other-directed, task-oriented behavior and fall and spring compliant behavior emerged as significant predictors of spring reading achievement for those pupils in the teacher directed settings. Different significant behavioral predictors were evident for those students in the program directed settings. Fall ratings of low peer dependency and high inner-directed, task oriented behaviors and spring ratings of low passive, withdrawn behavior and low inappropriately self-directed behavior were all significant spring reading achievement predictors for this group.

In the Teacher-Directed setting, pupils who were identified through initial observations as passive and withdrawn were found to be least likely as a group to gain in reading, while those who exhibited other-directed, task-oriented behavior and compliant behavior tended to gain more over the course of the year. In the Program-Directed sample, a decrease in low peer dependent behavior from fall to spring, and an increase in low passive, withdrawn behavior were found to be significantly related to growth in reading achievement.

The findings of the present review along with those of e.g., Kim, Anderson, and Bashaw (1968), Ross, Lacy and Parton (1965), and Kohn and Rosman, (1973) converge to define rather similar classroom behavioral

dimensions theoretically related to reading achievement. Lambert and Nicoll (1977) have defined these dimensions as 1) behaviors reflecting adaptation to the classroom environment e.g. following directions, volunteering, ability to work independently; 2) behaviors reflecting intrapersonal adjustment, e.g. withdrawal, attention; and 3) behaviors reflecting interpersonal adjustment, e.g. fighting, arguing.

Cobb (1970, 1972a) conceptualized the behaviors attending, compliance, volunteering, etc. as "academic survival skills" which were necessary, but not sufficient for successful academic functioning. They were not, he hypothesized, academic behaviors per se, but rather, the first components in a chain of correct academic responding; behaviors that all children, regardless of ability level, must first be taught in order to enable them to take advantage of the instructional opportunities in the classroom. Cobb and his associates have subsequently demonstrated the functional relationship of such behaviors to early reading development (Cobb and Hops, 1973; Hops and Cobb, 1974), wherein significant improvements were noted in both mathematics and reading achievement after initiation of "survival skill" training for underachieving first graders.

Particular moderating variables have been found to have a significant impact on measurement of the classroom behavior-reading achievement relationship at the elementary level. Differential behavioral predictors have been found between the sexes (Blom, et al. 1980; Cobb, 1970; Lahaderne, 1968; Samuels and Turnure, 1974; Swift and Spivack, 1967) and at different age or grade levels (Blom, et al., 1980; Coker and Lorentz, 1977; Feshbach, et al., 1977; McKinney, et al. 1975).

Intelligence has been identified as an important confounding factor because of its relatively high correlations with both reading achievement and many behavioral variables (Engin, 1975; Feshbach, et al., 1977; Harrison, 1975; Lahaderne, 1968; Meyers, et al., 1968; Miller, 1972).

Past evidence would suggest that sex of child, age or grade level of child, and intelligence are variables which should be either considered or controlled for in future analyses of the classroom behavior-reading achievement relationship. Further, as Cobb (1970) has suggested and demonstrated, the more precisely defined the classroom behavior is, the higher the relationship with achievement criteria. In addition, behavioral definition of student social-emotional status enhances further replication and clarification of research results, as well as providing useful information to educators in terms of defining non-intellectual variables which are potentially under the teacher's control.

Longitudinal Prediction of Reading Achievement

The problem of predicting reading achievement and the identification of children who are at high risk for school failure has been a concern of educators for many years. DeHirsch, Jansky, and Lankford's Predicting Reading Failure in the mid 1960's stimulated considerable interest in the development and use of various screening test batteries to predict reading failure. Efforts have most exclusively focused upon identification of "school readiness" skills in preschool and kindergarten age populations found to be significantly related to later reading success or failure. Significant emphasis has been placed upon the use of psychometric procedures which unfortunately have assessed a rather delimited set of linguistic and perceptual-motor skills (Adelman and Feshbach, 1971; Lambert and Nicoll, 1977). While most of the aforementioned

studies have yielded significant correlations between their predictors and subsequent achievement criteria, the relationships have generally been weak, particularly when and if they were subjected to cross-validation procedures. Further, varying degrees of predictive accuracy are reported (Cohen, 1963; DeHirsch, Jansky, and Lankford, 1966); Duffy, et al. 1972; Friedman, et al., 1980; Haring and Ridgeway, 1967; Klein, 1977; Lesser and Bridges, 1973; Meyers, et al., 1968; Satz and Friel, 1978; Stevenson, et al., 1976).

Among the more successful cognitive measures used to predict achievement have been those which have evaluated word and letter identification, visual and auditory discrimination, vocabulary and language development, and general intelligence. In their review, Stevenson, et al., (1976) suggested, however, that an understanding of the relative utility of the various measures have been difficult to attain due to the differing characteristics of the students sampled, the relatively small number of subjects per study, and the fact that prediction has rarely been over periods longer than one to two years. Further, few studies have attempted replication or cross-validation procedures to assess the empirical and practical utility and accuracy of their predictive instruments' (Meehl and Rosen, 1967; Satz and Fletcher, 1979).

Only in a minimal number of investigations have student behavioral characteristics been used in the development of instruments for longitudinal prediction of reading success or failure. Although such behavioral attributes of students have long been recognized as important factors in the development of reading competency, they have for the most part been excluded from school readiness predictive scales. Despite recognition of the potential importance of emotional/behavioral

functioning, investigators such as DeHirsch, et al., (1966) have systematically chosen to exclude such variables:

> We recognize that a variety of social, environmental, and psychological factors are significant in the acquisition of reading skills, and we concur with Fabian (1951), who maintains that learning to read requires the developmental timing and integration of both neurophysiological and psychological aspects of readiness. Nevertheless, we limited ourselves to the preschool child's perceptumotor and linguistic functioning because in this area we had found considerable deviation from the norm among children who subsequently failed in reading and spelling (DeHirsch, et al., 1966, p. 5).

Results of longitudinal studies which have used non-intellectual variables to predict reading achievement have suggested that student behavioral attributes may make a significant, unique contribution to the variance in later reading ability in the elementary as well as secondary grades, independent of various measures of cognitive ability, including measured intelligence.

Kellam, et al., (1975) for example, found that even when intelligence test scores and school achievement were included in a prediction battery, ratings of students "social adaptation status" in first grade made a significant, unique contribution to the prediction of later need for remedial help and grade retentions in the intervening elementary years.

The primary goal of Meyers, et al., (1968) study was to assess the predictability of fifth grade achievement and intelligence from a variety of measures of cognitive ability administered in kindergarten. These measures included tests of psychomotor coordination, perceptual speed, and linguistic and figural reasoning abilities. As an aside at the time of initial testing, students were rated by the examiners on ten different behavioral characteristics. Step-wise multiple regression

procedures were implemented and results suggested that a picture vocabulary measure appeared to be the single most valid predictor among the ability measures (r=.33-.50). The authors noted, however, that, "most striking in the results is the value of the behavior ratings in anticipating later achievement." Examiner's ratings of students' "attention" emerged as the single best predictor of fifth grade reading achievement, including both vocabulary and comprehension, as well as spelling achievement (r=.38-.43). Other behavioral indices which entered the prediction equations at significant levels included amount of motor activity, performance rate, effort displayed, and cooperation.

Lambert, et al., (1976) investigated the comparative ability of various "intellectual" versus "non-intellectual" indicators in predicting various aspects of effective high school functioning. As part of an earlier study data was collected on 300 second and fifth grade students which included teacher, peer, and self ratings on various behavioral dimensions, including compliance, ability to work independently, and interpersonal skills, grades, attendance records, and available health data. Fifty-six percent of the original sample were identified for the longitudinal follow-up. "Successful or unsuccessful high school status" was defined as a composite score based on such indicators as guidance referrals, referral to remedial reading classes, delinquency, participation in athletics, placement on the honor roll, and referral to gifted programs. A second set of criterion measures resulted from a factor analysis of such variables as counselor ratings of social and personality characteristics, standardized achievement and ability measures, GPA, and disciplinary referrals.

Multiple regression procedures were used to determine the relative degree to which status in high school could be predicted from cognitive and behavioral variables measured in elementary school. Teacher ratings of student behavior emerged as the single best predictor of successful and unsuccessful high school status (r=.42-.56). Although IQ entered significantly in the predictions, it contributed only minimally to the total variance after the teacher ratings were considered (r^2change=.01-.05). Unfortunately, the multiple R^2's, including both teacher ratings and IQ in predicting successful or unsuccessful adjustment, were relatively low (R^2change=.19-.36). It appears that these low correlations may have been an artifact of the relative lack of precision in the definition and measurement of the criterion. This point would appear valid in that significantly higher simple and multiple correlations were obtained between teacher's ratings and IQ and "high school scholarship," which was a composite factor score based on GPA, and standardized achievement and ability test scores. Again the teacher ratings of student behavior in elementary school proved to be the best predictor of later high school academic achievement (r=.47-.63). Intelligence was also highly related to high school academic achievement (r=.44-.58) and together with the teacher ratings accounted for approximately fifty percent of the variance in "high school scholarship."

One shortcoming of this study, which the authors recognized, was the fact that elementary achievement test scores were not incorporated into the prediction batteries:

> Overall, the best single predictor of successful status in high school was the set of teacher ratings items. While we did not analyze the efficiency of a battery which incorporated teacher rating items plus achievement scores and grades, it might have proved to

> be the set of predictors with greatest efficiency in estimating successful status and achievement in high school (Lambert, et al., 1976, p. 116).

Stevenson, et al., (1976a) further expounded on this notion as they suggested:

> Predicting reading achievement is a different problem before children have begun to read than after instruction in reading has been introduced. Before children learn to read, reading achievement must be predicted from tests of prereading skills and cognitive abilities. Once the child has entered elementary school, however, prediction can be made from the child's current level of achievement. For example, Henderson, Fay, Lindemann, and Clarkson (1973) gave over 700 7-year-olds a battery consisting of commonly used psychometric tests, including achievement tests. Scores on the test for reading achievement correlated .72 with reading achievement 1 year later. Adding scores from the remainder of the extensive battery raised the correlation only to .75. Thus, with children who already have begun to read, a battery of psychometric tests may add little to the predictive value of achievement tests alone. (p. 378).

Feshbach, Adelman, and Fuller (1977) further investigated the relative contribution of non-intellectual versus cognitive variables to subsequent reading achievement in the early elementary grades. Their results again suggested the importance and unique contribution of student behavior to the prediction of future reading attainment. The primary purpose of their study was to evaluate and contrast the predictive efficacy of the DeHirsch-Jansky Predictive Index, a measure of linguistic and perceptual-motor skills, and teacher's ratings of student cognitive and classroom behaviors as measured by the Student Rating Scale (Adelman and Feshbach, 1971). Their initial sample included 888 kindergarten age children in ten middle class Los Angeles schools. An additional 844 children were identified as a cohort sample for purposes of cross-validation of the findings. The authors suggested that the reason for limiting their population of interest to middle class children

was "to eliminate the powerful influence of social class variables on school performance so as to permit a clearer assessment of the role of individual predispositional factors." Of the original sample 403 or 45% were still available for testing at the end of third grade. This final sample included all students who had taken all three reading tests, the DeHirsch, and Student Rating Scale (SRS), and whose IQ's were 90 or above. Test scores from Cooperative Primary Reading Tests were obtained at the end of the first, second, and third grades and served as the primary predictive criterion.

Feshbach, et al. reported that for second grade reading achievement normalized SRS scores proved to be the best predictor, with an average correlation of .50. Full scale IQ and DeHirsch scores were similarly high related to second grade achievement (r=.44 and .42 respectively). Similar correlations were obtained for the third grade reading achievement criterion. Using multiple regression procedures it was found that with IQ partialed out, the correlation of SRS with third grade reading was .33. Highest regression weights were obtained for IQ and SRS with the multiple R based on these two variables being .57. Cross-validation with the second cohort group collaborated these results as comparable relationships were found (multiple R=.50). Multiple Rs for second grade achievement were similarly consistent for the initial and cohort groups (r=.57 and .61 respectively).

Stevenson, et al. (1976a) further added credence to the findings of the ability of teacher ratings of pupil behavior and cognitive skills in kindergarten to predict future reading achievement in the first through third grades. In contrast to the findings of Feshbach, et al. (1977), Stevenson and his colleagues indicated that teacher ratings

were not found to be as predictive of later achievement as were scores on various cognitive and psychometric tasks. It would appear that though the scales and tasks differed in the two studies, the differences in results arose, not because the correlations of teacher ratings were lower in the Stevenson, et al. study, but because the correlations of the test scores were significantly higher (Multiple R's=.56-.77). This appears attributable to the relative restrictive range or homogeneity of the Feshbach, et al. sample as compared to that of Stevenson, et al.

In a subsequent follow-up study Stevenson, et al. (1976b) evaluated more specifically the predictive value of the teacher ratings obtained in the original study independent of the cognitive and other psychometric variables. A total of 217 children were initially evaluated and rated by teachers on 13 scales which represented such attributes as intellectually independent, hardworking, reflective, follows instruction, retains information, vocabulary, attention, and social acceptance. Variables for the scales were selected on the basis of "extensive discussions about characteristics of children that appear to be important for successful performance in school." The authors' report of items included in the scale suggest high face validity, however, no measures of scale reliability were reported. Further, interpretation of their results appears somewhat confounded by the fact that various formats were used in obtaining the scale scores.

As in the aforementioned Stevenson, et al. studies, stepwise multiple regression procedures were implemented. Teachers' ratings of students' ability to follow instructions, vocabulary development, ability to retain information, intellectual reflectiveness, intellectual independence, and ease of understanding concepts (Effective Learning) all

entered prediction equations at significant levels at grades one through three. A simple sum was computed for the four most frequently selected variables (Effective Learning, Vocabulary, Following Instructions, and Retaining Information) and this sum subsequently correlated with reading achievement scores. These were as follows: first grade achievement (r=.66); second grade achievement (r=.65); and third grade achievement (r=.55). The values of these correlations were found to be essentially equal to those obtained by the stepwise regression procedures. The correlations obtained from the sum of these four ratings were further found to be highly commensurate with correlations obtained from the sum of the four most predictive cognitive and psychometric tasks in the original study (Stevenson, et al., 1976a).

Of the longitudinal studies which have investigated the utility of a variety of screening indices in predicting future reading achievement, some have further attempted to ascertain the accuracy and practical value of their predictive instruments. Satz and Fletcher (1979) address the relative problems of determining predictive outcomes for students as well as the implications for such procedures upon actual programmatic or placement decisions:

> Because of the current interest in early screening, attempts are being made to identify the high-risk child during preschool and to place him in intervention programs before his problems become more severe or refractory to remediation. However lofty these ideals may be, they portend unnecessary cost to the schools and risk to the individual child if the predictive utility of the early detection device is inadequately assessed. A true false positive and false negative error rate is frequently masked by obvious confoundings from misclassification of high-risk Ss (false negatives) or low-risk Ss (false positives), making early detection simply not possible (p. 65).

Those studies which have attempted to determine the relative predictive accuracy of their instruments in terms of correct prognostic classification have reported varying degrees of success. It would appear that much of the variation in predictive accuracy in these studies may be attributable to 1) the degree of the relationship between predictor and criterion variables, 2) the nature of the criterion variable e.g. teacher report or test score, 3) the variability and "arbitrariness" in determination of criterion cutting scores, and 4) the general methodological means by which predictive accuracy was determined. Such variation in methodology will become evident as this review proceeds.

Lesser and Bridges (1973) investigated the predictive effectiveness of the Metropolitan Readiness Test, the Lee-Clark Readiness Test, and the California Test of Mental Maturity in determining whether or not a student would eventually manifest a "learning problem." Students were initially evaluated on the predictor variables early in their first grade year. Criterion measure of achievement, the California Achievement Test (CAT) and a Teacher Rating (TR) of overall performance, were obtained at the end of the first and second grades. Because of the relatively high correlation between achievement test scores and teacher ratings (r=.77) the investigators decided on a "combined criterion." In order to combine the CAT score and the TR, each score was divided into categories of "no learning problem" (NLP) and "learning problem" (LP). Teachers determined that a CAT score of 134 or more, which corresponds to a grade placement of 1.5 or higher, would be considered to be reflective of NLP, whereas a score of 133 or less indicated LP. In terms of the TR, either of the two highest ratings ("average or better"

and "marginal") was considered NLP, while the third rating ("clearly immature") indicated LP. On the combined criterion all subjects who had been classified as NLP on both the CAT and the TR were classified NLP. All subjects who were classified as LP on either of the two single criteria were categorized as LP. On the combined criterion, 148 children (51 percent) were classified NLP and 145 children (49 percent) were classified LP.

In order to test and compare the relative predictive effectiveness of the Metropolitan, Lee-Clark, and California Test of Mental Maturity (CTMM), the number of correct predictions made with each of these instruments were calculated. For these analyses the authors suggested that "cutting scores were used which minimized incorrect predictions of LP." These scores which were employed as the upper limits of the prediction of LP, were 35 on the Metropolitan, 42 on the Lee-Clark, and 34 on the CTMM. The investigators did not indicate the rationale for how these scores were determined.

Results of the study suggested that the Metropolitan was the most effective test in predicting first grade performance, as it "correctly predicted a total of 18 to 19 more children than did the other two tests." The percent of students correctly identified as LP of NLP were reported for each of the predictive measures. The Metropolitan correctly predicted 86% of the LP students and 87% of the NLP students for an overall "hit rate" of 86%. The Lee-Clark correctly predicted 82% of the LP students and 78% of the NLP students for an overall hit rate of 80%. Finally, the CTMM was reported to correctly identify 82% of the LP students, 79% of the NLP students, and 80% overall. Of those children available for testing at the end of second grade, 119 were found to have

problems and 77 performed adequately. Using the MET cutting scores of 36/35 employed in the prediction of first grade performance, 74 (91%) of the children predicted LP were classified LP and 70 (61%) of those predicted NLP did, in fact, perform adequately.

The observed primary limitation of this study is that it did not address the practical application, as well as limitations of using such instruments as screening measures for making programming decisions for individual students from the predictive information. The investigators did not address the issue of those students who were incorrectly classified, that is those students who were predicted to be NLP, but indeed were LP - 39% in their sample for the second grade predictions. As such, the practical utility of such measures, e.g. Metropolitan, Lee-Clark, or CTMM, may indeed be suspect for purposes of individual decision-making.

Feshbach, Adelman, and Fuller (1974) examined the relative efficiency of the DeHirsch Predictive Index versus teacher ratings of students' skills and behaviors in predicting "high or low-risk" students as measured by a determined cut-off score on the Gates-MacGinitie Reading Test. Students were initially evaluated on the DeHirsch and Student Rating Scale (SRS) while in kindergarten, and reading achievement data was obtained at the end of the students' first grade year.

A step-wise multiple regression analysis was carried out such that the independent contribution of the predictors could be determined. The SRS emerged consistently as the best predictor (r=.57). Further, addressing the issue of predictive accuracy, Feshbach, et al. suggested that:

> . . . for the purpose of early intervention, a more relevant datum than overall correlation is the efficiency in predicting those children who, by virtue of their falling below a specific criterion, can be considered reading failures (p. 53).

A grade placement score of 1.7 and an IQ greater than 90 were selected for defining the criterion of good versus poor reading performance. For the SRS an arbitrary Z-score was selected as the cutoff point for purposes of prediction from the teacher ratings, whereas for the DeHirsch, a previously established cutoff criterion of less than five of nine items passed was used for predicting future performance. Relative overall accuracy of prediction was comparable for the DeHirsch, the SRS, and a combination of the measures, ranging from 72-77% correct predictions. The authors further reported that the DeHirsch generated twice as many false positives as the SRS (39% versus 17% respectfully), and that the proportion of false negative errors was relatively small for both measures.

Satz and Fletcher (1979) point out that Feshbach, et al. (1974), however, inadvertently reversed the rows and columns of their detection tables and also reversed their detection rates such that their positive signs (valid and false) each summed to unity. Satz and Fletcher subsequently corrected these detection rates and found that the SRS had a 30% valid-positive rate, a 70% false-negative rate, a 3% false-positive rate, and a 97% valid-negative rate. The overall hit rate for the SRS was 77%. By comparison, the DeHirsch had a 26% valid-positive rate, a 74% false-negative rate, a 7% false-positive rate, and a 93% valid-negative rate. The overall hit rate was 73%.

The major errors in the original Feshbach, et al. (1974) study were made in terms of the relative percentages and not in terms of the absolute values of the false-negative and false-positive errors. The most significant inaccurate conclusion in the original study was that the false-negative rate was rather small for both measures (24-25%). The corrected rates revealed that the false-negative rates were extremely high (70-74%). Satz and Fletcher further point out that both the SRS

and DeHirsch failed to detect over 70% of the high risk children, "a finding that may have been overlooked given the incorrect percentages reported by the authors." They suggested that if an intervention program were to be implemented based on these percentages, many potentially high-risk children would have been excluded from the prevention program.

The results of Stevenson, et al. (1976a) suggest the relative drop in predictive accuracy over periods of more than one or two years. Their previously identified optimal predictor variables generated through multiple regression procedures in kindergarten sample were evaluated to see how well they identified students who actually had later reading difficulties, as indicated by their scoring in the lowest 10 or 25% of the third grade. The findings indicated overall low predictive accuracy figures relative to actual student performance. Only one third of the children in the lowest 10% in third grade reading also scored in the lowest 10% on the prekindergarten screening tasks. Again false-negative predictions were inordinately high. Stevenson, et al., concluded that in order to identify more than half the children who had serious difficulty in third grade (lowest 10%) from the prekindergarten predictors, it would have been necessary to select 25% of the children on the basis of their performance on the prekindergarten tasks. Similarly, to identify students having at least moderate difficulty in reading in third grade (lowest 25%), it would have been necessary to select 40% of the children prior to kindergarten.

Stevenson, et al., further found only moderate success in terms of kindergarten teachers' predictions of students' later reading difficulty. Of the twelve children named by their kindergarten teachers and for whom third grade data were available, only three were in the lowest 10% and six were in the lowest 25%.

Satz and Friel (1974) investigated the predictive accuracy of twenty-two various developmental and neurological tests initially administered in kindergarten in determining subsequent first grade achievement. The criterion measure of reading achievement in this study was based upon a ten item scale of reading level which was assessed by the teacher for each child at the end of the first grade. The scale ranged from No Readiness (score=0) to Advanced Third Reader (score=9), with Primer and First Reader scored as 4 and 5, respectively. Those whose reading level was rated as below Primer (levels 0-3) were assigned to the High Risk group; those whose reading was judged to be at the Primer level or above (levels 4-9) comprised the Low Risk group. The two criterion groups (High and Low Risk) were further subdivided into an extreme and conservative reading subgroup. The two High Risk subgroups were Severe (levels 0-2) and Mild (level 3) and the two Low Risk groups were Average (levels 4-5) and Superior (levels 6-9).

A unique methodological aspect of this study was the use of discriminant function analysis to ascertain the most parsimonious combination of predictor variables that would maximize the differences between the identified groups (high and low risk). A composite discriminant predictor score (Z_i) was computed for each student, and these scores were then averaged for each group and tested for significance. Results indicated that 78% of the High Risk group were correctly classified. The false-negative rate was 22%. Similarly 86% of the Low Risk group were correctly classified with a false-positive rate of 14%. The overall hit rate was 84%.

The predictive classifications were further examined within criterion groups in order to determine whether accuracy varied as a function

of extreme and conservative categories in the High and Low Risk groups. Results demonstrated that accuracy was higher in the extreme categories. All of the severely disabled readers were correctly classified and 95% of the superior readers were correctly classified. Virtually all of the misclassification errors were confined to the more conservative categories, Mild High Rish and Average Low Risk (all of the false-positives and 95% of the false-negatives).

Satz and Friel finally addressed the notion of the predictive utility of their battery:

> The concept of predictive utility implies that the detection signs (e.g., valid positives and negatives) should be evaluated within the broader context of the prevalence estimates of the event to be predicted. . . in this case this refers to the incidence or base rate of reading disability (Satz and Fletcher, 1979, p. 66).

Because of the relatively high false-positive rate obtained with the original cutoff criterion, Satz and Friel elevated the cutoff criterion in order to reduce the likelihood of misclassifying a normal child as High Risk and instituting treatment. By optimizing the valid and false positive rates, the conditional probablities were subsequently computed for the differential composite signs (+ and -). Given an observed base rate of reading disability of 15% in the sample, the probability that a student would be high risk, given a positive composite score, was .72. Conversely, the probability that a child would be Low Risk, given a negative composite score, was .94.

In a subsequent follow-up study Satz and Friel (1978) investigated whether their abbreviated test battery given at the beginning of kindergarten could predict reading group membership at the end of second grade on a new mixed sample of kindergarten children (males, females, black, white) based on the discriminant lambda weights derived from the

two standardization populations of white males.

At the end of second grade individual reading measures were obtained on 80% of the original sample. On the basis of those measures, students were divided into Severe, Mild, Average, and Superior reading ability groups. The Severe group was reading approximately 1.25 years behind grade level, whereas the Superior group was reading almost 2.5 years beyond grade level expectancy.

Results indicated that the composite scores from the tests in kindergarten correctly predicted 100% of the Severe and Superior groups, while misclassifying 19% of the Mild group and 37% of the Average group. The overall hit rate on this cross-validation sample was 75%. Again conditional probabilities for each of the composite test signs were calculated. The probability that a child would be high risk, given a severe high risk composite score in kindergarten, was .72. This composite score, which was the basis for treatment in this sample, managed to screen 90% of the severe cases and 69% of the mild cases. A much lower predictive outcome was revealed for the mild high risk composite score, .27. In contrast predictions were much higher for both low risk composite scores, .92 and .96.

Like Satz and his colleagues, Friedman, et al. (1980) found that despite relative overall high predictive accuracy of the latter screening instruments, the rates of false-positive predictions appeared relatively high and posed significant question as to the practical utility of their prediction battery. Friedman, et al. (1980) assembled a multifactional battery composed of the Star Picture Vocabulary Test, the Goodenough-Harris Human Figure Drawing Test, the Berry-Butkenica Development Test of Visual-Motor Integration (VMI), and the Personal-Social Responsiveness subtest of

the Caldwell Preschool Inventory. This battery was initially administered to a kindergarten sample. Primary criterion variables included reading achievement test scores on the Cooperative Primary Tests, as well as teacher questionnaire data on overall academic achievement and reading proficiency. Findings were reported for three population groups for whom teacher ratings were obtained, 1) all students whose school performance was assessed at age seven years, 2) 1,251 children evaluated at age nine years, and 3) 600 children for whom ratings were available at both seven and nine years.

For the first group, (those students assessed at age seven years) the Caldwell Test was the best predictor, but had significant value for girls only. However, of the girls who failed this test, 59% did not have a problem at seven years. Thus, a high false-positive rate occurred, even with markedly significant association. Combining tests did not significantly increase the predictive value of the Caldwell Test alone. In the second group (those students assessed at age nine years), for all children and for each sex separately, the VMI and the Caldwell were the best predictors with the VMI having slightly higher efficiency than the Caldwell. False-positive rates were again high-62% or more. For the third group, the VMI test repeated at age seven years predicted academic achievement and reading at age nine years at a significant level for both girls and boys. The best predictor, however, was the teacher's report of academic performance at age seven years; the false-positive rate was again high at 42%.

Combining the academic performance at age seven years with the seven-year VMI yielded an increased accuracy of prediction. Using these combined criteria, 89% of the children were correctly classified with regard to success or failure at age nine. However, this "hit rate" included both the correct positive and the correct negative predictions, which as have

been discussed are usually not of equal practical importance. For example, of those children who failed the combined criteria, 65% had problems at age nine. This 65% is high when contrasted with 11% of the school failures in the population of nine-year-olds. However, it must be kept in mind that the remaining 35% who failed both the VMI and the teacher report at age seven, did not have school problems at age nine and would therefore have been mislabled. If the consequences of misclassifying these false-positive children into a high risk group were significantly detrimental, this rate of 35% might be considered excessive, even though there is strong statistical significance.

Sex Differences in Reading Achievement and Classroom Behavior

Boys comprise the vast majority of educational problems; many estimates put the figure at near eighty percent. Although there appears to be no sex differences in terms of general intelligence, girls as a group, especially in the early elementary years, achieve significantly higher than boys in the areas of reading, spelling, and writing. Estimates of the number of boys who experience difficulty with the reading process at some time during their academic experience are estimated to range from seventy-give to ninety percent (Gentile and McMillan, 1976; Stanchfield, 1973). Blom (1970) noted that sex ratios in studies of reading disability range from 1.3:1 to 15:1 with boys representing the higher figures. Entwisle (1971) estimated that over ninety percent of referrals to reading clinics were males. Furthermore, the ratio of boys to girls referred to learning disabioity classrooms appears to be fairly consistent across studies, that is approximately 4:1 (Bannatyne, 1971; Bentzen, 1966; Naiden, 1976).

Along with difficulties in reading, boys as a group also experience a number of potentially related problems during the course of the elementary years. Boys are consistently rated by teachers as exhibiting significantly higher rates of disturbing classroom behaviors, e.g. aggressive, defiance, impulsiveness, hyperactivity, low need for achievement, and inattentiveness (Goyette, Conners, and Ulrich, 1978; Miller, 1975; Peterson, 1961; Spivak and Swift, 1967). On attitude surveys boys have been found to be more negative towards school in general (Berk, Rose, and Stewart, 1970) and toward reading in particular (Neale, Gill, and Tismer, 1970). Bentzen (1966) reported that while boys were referred three times more often than girls for academic related problems in the first grade, they were referred eleven times more often for reasons of "social and emotional immaturity." Morse, Cutler, and Fink (1954) found the ratio of boys to girls in classrooms for the emotionally disturbed to be five to one. A review of relevant studies of the relationship between classroom behavior and academic achievement of elementary school children suggest that sex of student may be relevant moderating variable in this relationship. Differences with respect to reading achievement in the elementary grades between boys and girls (Maccoby and Jacklin, 1974; Thompson, 1975) and with respect to classroom behavior (Rubin and Balow, 1978; Werry and Quay, 1971) have been established. It may, however, be the case that a pattern of classroom behaviors may be predictive of reading achievement for girls but not boys, or it may be that the direction of the relationship between behaviors and subsequent achievement is different for male and female elementary pupils.

While Lahaderne (1968) and Samuels and Turnure (1974) found that sex of pupil did not have an effect on the behavior-achievement relation, a number of other investigators have found support for such a hypothesis.

Cobb (1970) for example found a higher correlation between attention, compliance, and volunteering and achievement for boys than girls, and he further reported that different behaviors entered regression equations for the two sexes predicting first grade achievement. Blom, et al. (1980) found a significant sex by reading ability interaction, wherein poor reading males were rated by teachers as exhibiting much more acting-out, aggressive behavior than their good reading male peers. On the other hand, post-hoc comparisons of the aggression scores in the female sample failed to reveal a significant difference between good and poor readers, although mean differences indicated good reading girls were rated higher in aggressive behavior than their poor reading female counterparts.

Feshbach, et al. (1977) found several interesting differences in a longitudinal study of the relationship between teacher ratings of classroom behavior and reading achievement for first through third grade boys and girls. They found that the overall and attention-behavior control scores of the Student Rating Scale (Feshbach and Adelman, 1971) correlations for girls were higher than those of boys. They suggested that:

> Motivational variables may be more important for girls in the early primary grades, thus possibly accounting for the lesser significance of IQ (as compared to boys). At the same time, kindergarten behavior factors, which reflect poorer performance by males than females, may decline in importance for males with increasing maturity (p. 303).

Analysis of behavior-achievement data generated during the course of the investigation of Blom, et al. (1980) indicated differences in the early versus later elementary years on the relationship between various behavioral dimensions and reading achievement between the sexes. For example, good reading females in the lower elementary grades were rated higher on an aggression scale than their upper elementary peers. In addition, poor reading males were rated higher in aggression in the lower

elementary grades were rated higher on an aggression scale than their upper elementary peers. In addition, poor reading males were rated higher in aggression in the lower elementary grades than their upper elementary counterparts.

Sex Differences in Reading Achievement in the Elementary Grades

In this section the evidence of sex differences in reading achievement in the elementary grades is reviewed. To summarize, there appears to be a developmental trend regarding differences between the sexes in reading achievement, wherein, a significantly larger proportion of boys than girls make a slow beginning at learning to read, but during the later elementary years, grades four through eight, significant population differences begin to disappear.

Sex differences in reading ability have been a source of concern to educators, researchers, and school administrators for many years. As early as 1909, Ayres in his book, Laggards in Our Schools, concluded: "Our schools as they now exist are better fitted to the needs and natures of the girl than the boy pupils" (p. 155). Subsequent empirical studies have confirmed the observation that boys in American elementary schools are far more likely than girls to have serious reading difficulties.

Examination of investigations from several English-speaking countries, primarily the United States and Great Britain, indicates that the trend of sex differences in mean reading achievement scores for children approximately ten years and younger is different from that of older elementary students; girls as a group generally score significantly higher than boys in the early grades. Mean scores in both reading comprehension and reading vocabulary were found to be significantly lower for boys than girls by Carroll (1948) for six to seven year olds, and by Preston (1962)

and Rutter, et al. (1970) for nine year olds. In all cases the mean achievement levels of the boys were significantly lower than that of the girls. On standardized measures of word recognition, boys again scored significantly lower than girls in studies reported by Southgate (1959) for six to seven year olds, and by Kellmer Pringle, et al. (1966) in a nationwide sample of seven year olds in England.

Numerous studies of first grade children have confirmed the reading achievement superiority of girls during the early stages of reading instruction. Konski (1951) evaluated the performance of boys and girls on tests of reading readiness at the beginning of the academic year, and subsequent reading achievement at the end of the year. Results indicated no significant differences between the sexes on the readiness measures, however girls scored significantly higher on the reading achievement measures at the end of the year.

Nila (1953) similarly tested three hundred first graders during the first weeks of school on a number of individual and group administered reading readiness tests. Again it was reported that boys as a group and girls as a group were comparable in readiness to read. The students were again tested at the end of the year. Seventy-two pupils were designated as "reading failures," and of these two thirds were boys.

During the course of the 1962-63 school year in the Los Angeles City Schools, Stanchfield (1965) initiated a study of the effect of homogeneous grouping of first grade pupils by sex upon subsequent reading achievement. Over five hundred students took part in the study. The analysis of reading achievement and reading growth indicated no significant differences between boys taught in all-male groups and those taught in the sexually heterogeneous groups. The girls on the whole, irregardless of group,

however, scored significantly higher than the boys at the year's end.

During the school year 1964-65, twenty-seven first grade reading studies were facilitated by grants from the U.S. Office of Education. A large percentage of these studies compared, as did Stanchfield, the relative efficiency of different methodologies and materials in the first grade. Some of these studies did not directly compare scores of boys and girls, while others sought to determine if one particular instructional approach was more effective with either boys or girls than were other approaches. A number of these studies do provide data on sex differences in reading achievement in grade one. Of significance is the fact that each of these studies utilized a number of common pre- and post-treatment measures of achievement, primarily the Stanford Achievement Primary Battery. Results from these studies which compared the relative achievement of boys and girls are cited.

Tanyzer and Alpert (1966) compared the mean scores achieved through the use of three basal approaches (Lippincott; ITA; and Scott-Forsman) and found that in each of the three basal systems, the girls achieved significantly higher mean scores than the boys with significant differences favoring the girls on the Spelling and Paragraph Meaning subtests. Hahn (1965) reported comparative data only for the Word Study subtest, where girls scored significantly higher than boys. Schneyer, et al. (1966) noted significant differences favoring girls on Paragraph Reading, Spelling, and Word Study Skills.

Spache (1965) and Fry (1966) both noted significant differences favoring girls. Fry also noted that none of the specific methods investigated (ITA; Diacritical Marking; and regular basal materials) proved to be more advantageous for boys as a group.

Other studies of second and third graders have confirmed early female superiority in reading achievement in the elementary grades. Hughes (1953) evaluated over six hundred students in the metropolitan Chicago area in grades two through eight. Utilizing the total comprehension score of the Chicago Reading Tests, she found the greatest mean differences between boys and girls in her third grade sample. Mean grade level scores were higher for girls at all grade levels, except for fifth grade, where factors of chronological age and mental ability significantly favored the boys. Significant statistical differences favoring the girls were found in the third and fourth grade samples.

Blom, et al. (1980) similarly found significant reading achievement differences favoring girls in both the second and third grades utilizing a combined Vocabulary and Comprehension score from the Gates-McGinitie Reading Achievement Test. Consistent with the results of Hughes (1953), significant differences diminished in the upper elementary grades, four through six.

Measures of reading attainment discussed to this point in the review have been exclusively group administered, standardized tests. A few reports of sex differences in the early elementary grades on individually administered tests or oral reading are, however, available. The findings of Rogers (1971) and Prawat and Jarvis (1980) appear to confirm the superiority of girls in the early elementary school grades on such reading tasks. Rogers (1971) found in a sample of 500 students aged five and a half to six years that boys as a group achieved significantly lower on an individually administered oral reading task. Prawat and Jarvis (1980) in a study of first through third grade students in a large midwestern urban school setting found girls to score significantly higher on the

Slossen Oral Reading Test at all three grade levels.

Anderson, Hughes, and Dixon (1956) investigated individual differences in reading development in elementary age pupils at the University of Michigan Laboratory School. Significant sex differences were noted in the chronological age of pupils of learning to read. The students were judged to have "learned to read" when they achieved a reading age of eighty-four months on the Gates Primary Reading Tests. Their primary findings were that:

1. Age of learning to read was widely distributed despite the fact that most of the children in the sample were of superior intelligence.

2. Girls tended to read earlier than boys and there were fewer extreme delays in learning to read among girls--a large percentage of boys actually learned to read at second grade rather than first. The average age of learning to read was 83.1 months for girls and 89.6 months for boys.

3. Age of learning to read was significantly corrolated with intelligence quotients recorded in the first grade (.57 for girls and .54 for boys).

4. Age of learning to read was significantly corrolated with subsequent reading achievement in the sixth grade (.67 for girls and .65 for boys).

In a second study, Anderson, Hughes, and Dixon (1957) found that after children reached the reading age of eighty-four months on the Gates Primary Reading Test, there was no difference between boys' and girls' subsequent growth rates in reading. Sex differences in the age of learning to read tended to disappear among children of high intelligence (above 100), whereas children of lower intelligence (less than 100) tended to

start later, and proceed generally slower than the higher IQ group. A high degree of variability in the rate of reading development was, however, observed in the lower IQ group. Sex differences were further most prominent in the lower IQ group with the boys tending to start later than the girls.

For children in the fourth through eighth grades the consensus of the literature indicates generally nonsignificant differences between the sexes in reading achievement (Thompson, 1975). Divergent findings tend to support the notion of female superiority throughout the elementary grades.

Numerous studies have indicated no significant differences in reading achievement between the sexes on group administered measures in the upper elementary grades (Blom, et al. 1980; Commins, 1928; Heilman, 1933; Heilman, 1961; Horton, 1973; Rutter, et al., 1970; and Wozencraft, 1963). Neale (1958) and Vernon (1938) furthermore found no significant differences between the sexes in the ten to fifteen year age group in individually administered tests of oral reading, as did Prawat and Jarvis (1980) with fourth through sixth graders.

Stroud and Lindquist (1942) found the mean achievement levels in both reading comprehension and reading vocabulary to be significantly lower for boys in the third and fourth grades in their cross-sectional study of third through twelfth grade students in Iowa. They further reported nonsignificant mean differences between the sexes at all other grade levels favoring the girls in reading vocabulary and comprehension, language, and word study.

Assessment data from the State of California 1975-76 testing program of second, third, and sixth grade students substantiated the finding of

female superiority throughout the elementary grades in reading, written expression, and spelling (Gehringer and Barnes, 1978). Results of recent Michigan Educational Assessment data further substantiate the reading attainment superiority of girls from the fourth through tenth grades:

> Staff found that on each of the fourth grade reading objectives, the attainment rates for girls were higher than that for boys by an average of 9.7% in 1979-80, 8.8% in 1978-79, and 8.5% in 1977-78. . . Further the assessment data show that this basic pattern holds true for seventh and tenth graders with boys closing the gap slightly at the higher grades (Katz, 1980).

Gates (1961) sampled over thirteen thousand pupils in grades two through eight to determine sex differences in reading ability. A total of twelve school systems in ten states participated in the study. Boys and girls were compared at each of the seven grades on reading speed, vocabulary, and comprehension. In each of the twenty-one comparisons, girls had higher mean raw scores with most of the mean differences significant. Preston (1962) obtained similar results in reading comprehension in his American sample of eleven year olds.

To summarize, there appears to be a developmental trend regarding differences between the sexes in reading achievement, wherein, a significantly larger proportion of boys than girls make a slow beginning at learning to read, but during the later elementary years, grades four through eight, significant population differences begin to disappear.

Sex Differences in Behavior

Elementary age boys have been shown to exhibit significantly more achievement-inhibiting, as well as disturbing classroom behaviors as girls in both normal and special class populations of students (Rubin and Balow, 1978; Spivak and Swift, 1966; Werry and Quay, 1971). Whether one looks at mental retardation (Lehrke, 1978), reading difficulty (Spache, 1976;

Rutter and Yule, 1977), hyperactivity (Cantwell, 1977), aggression (Maccoby and Jacklin, 1974; Miller, 1972), or simply low grades (McCandless, et al., 1972), boys predominate in frequency and intensity of the problem.

The focus of this review is upon those differences between the sexes that have been reported in regard to the prevalence of behavioral adjustment difficulties of elementary age children. It is beyond the scope of this review to address the possible causes for the observed differences, and the reader is referred to Eme's (1979) article for a more indepth examination of the topic.

In an investigation of nearly 1000 kindergarten through third grade children, Rubin and Balow (1971) found that educationally defined behavior problems (i.e., inability to meet school demands), were present in 41% of their sample. In all areas in which results were differentially reported by sex, boys consistently outnumbered girls, e.g. grade retention, attitude and behavior adjustment difficulties, and special class placement. In their 1978 longitudinal examination of teacher identified behavior problems, Rubin and Balow found the proportion of boys identified as behavior problems remained relatively constant at 35% over the seven year follow-up period from kindergarten through sixth grade for the same sample. However, considerable variability in the proportion of behavior problems identified among the girls was evidenced, ranging from a high of 26.4% at grade 1 to a low of 12.8% at grade 6. In any single year, approximately 69% to 76% of all the subjects in the study were classified as no problem. One of the more interesting findings of the study was that more than half (59%) of the 1527 studnets who had received three or more teacher ratings, were classified as a behavior problem at least once. Among the subjects who received six teacher ratings, 60% (68% of the boys and 51% of the

girls) were considered a behavior problem by at least one teacher. Despite the relative lack of precision in definition of a behavior problem in their study, Rubin and Balow's (1978) findings do indicate that boys are indeed rated by teachers as exhibiting more adjustment difficulties than girls, consistent with Brophy and Good's (1970) observation that, "boys are generally more salient in the teacher's perceptual field than girls."

Probably the most unequivocal sex difference in the literature is the male preponderance of aggressive behavior and generally higher levels of activity (Block, 1976; Eme, 1979; Maccoby and Jacklin, 1974). Despite relative developmental changes on these dimensions for both sexes, boys are rated by both parents and teachers as exhibiting significantly higher levels of aggression and "overactivity" as compared to girls.

Achenbach and Edelbrock (1981) in their investigation of the prevalence of behavior problems among normal school age children interviewed nearly 1500 parents and solicited their responses to the 112 item Child Behavior Checklist. ANCOVA procedures indicated significant gender differences on 51 of the items. Of the 51 main effects, 26 reflected higher scores for girls and 25 higher scores for boys. Nineteen of the 25 items on which boys scored higher were found to be primarily associated with what the authors conceptualized as "a broad band externalizing syndrome," which included such behaviors as bragging, hyperactive, cruelty to others, destructiveness, disobedience, fighting, impulsivity, showing off, temper tamtrums, and swearing. By contrast, only one of the items (screams a lot) on which girls scored higher was primarily associated with the externalizing syndrome; 10 items were primarily associated with the internalizing syndrome, and the remaining 15 items were not associated

with either of the syndromes. Girls scored significantly higher than boys on such behaviors as being too dependent, lonely, crying a lot, feeling unloved, nightmares, headaches, self-consciousness, shyness, moodiness, and worrying. These identified behaviors for boys and girls appear highly consistent with those of MacFarlane, et al., (1954) who found that boys exhibited significantly higher levels of overactivity, demand for attention, temper tantrums, and lying, while problems for girls centered more commonly around excessive reserve, shyness, and specific fears. Despite the clear differences between the sexes on these individual items the mean total behavior problem scores for boys and girls were virtually identical in the Achenbach and Edelbrock study and the authors suggested that:

> The much higher mental health and special education referral rates for boys than girls throughout childhood and their higher court referral rates in adolescence may result more from the conflicts between their behavior and official norms than from quantitative differences in problems (p. 51).

Werry and Quay (1971) again studied the prevalence of behavior problems in a nonclinical sample of 1700 students aged five to eight. Their results indicated that while symptoms of pyschopathology were quite prevalent overall in this age group, boys tended to exhibit more acting out or disruptive behaviors than girls. Forty-nine percent of the boys in the sample were rated as restless, 46% disruptive, 43% short attention span, 43% inattentive, and 48% distractible. "Distructiveness" was present in 16% of the boys and 5% of the girls, while "hyperactivity" was present in 30% of the boys and 14% of the girls. Werry and Quay concluded that there was little doubt that not only did boys have more psychopathological symptoms but the "connotative sense of most of the symptoms more common in boys represents 'badness'" (p. 142). Similar sentiments were expressed by Schlosser and Algozzine (1979) who found that those behaviors which tend

to be more characteristic of boys, e.g. acting-out, aggressive, immature, are perceived by teachers as more bothersome than those behaviors characteristic of girls, e.g. withdrawn, self-conscious.

Lapouse and Monk (1964) in a later analysis of their examination of maladjustment among a non-clinical sample of six to twelve year old children (1958), similarly found boys as a group to be rated significantly higher than girls by parents as exhibiting more extreme deviations on such behaviors as bedwetting, masturbation, overactivity, temper loss, and teacher complaints. Replication of the study with a rural elementary sample indicated similar results (Schultz, et al., 1974).

Spivack and Swift (1966) in their initial standardization of the Devereaux Elementary Behavior Rating Scales found significant sex differences on eight of the twelve factors of the teacher rating scale. Boys were found to score higher on those dimensions that were detrimental to achievement than girls, whether in special or regular class settings. Boys scored significantly higher on those dimensions defined as Classroom Disturbance, Slowness in Work, External Reliance, Need Achievement Recognition, Impatience, and Disrespect-Defiance. Girls scored significantly higher on those dimensions defined as Need for Closeness with the Teacher and Creative-Initiative (Verbal).

Similar differences between boys and girls were identified by Goyette, et al., (1978) in the restandardization of the Conners Parent and Teacher Ratings Scales. In a sample of 590 elementary age students, boys were rated by their parents as exhibiting more difficulties on those factors defined as Conduct Problem (e.g. destructive, quarrelsome, fights), Learning Problem (e.g. difficulty learning, distractible, easily frustrated, immature) and Hyperactivity (e.g. restless, disturbs others,

excitable, impulsive).

In their four year longitudinal study of the ability of teacher ratings of students' abilities, classroom skills, and personal-social skills to predict future reading achievement, Stevenson, et al. (1976a) found that girls were rated higher on various achievement-enhancing behavior dimensions in all comparisons. Girls were given significantly higher ratings in Overall Adjustment and Persistence, while mean ratings were also higher in Effective Learning, Hardworking, Retaining Information, Reflective, Attention, and significantly lower in Activity Level.

Validity of Teacher Ratings of Pupil Behavior

Because the judgments of classroom teachers have been increasingly used as a first indicator of behavioral problems of students and may be a primary referral source to special education or remedial services, a number of researchers have attempted to systemize teachers' observations of pupil behavior in order to achieve the measurement characteristics of standardized instruments. This has resulted in the construction of numerous inventories of child behavior in which norms referenced procedures have been used to establish base rates of problem behavior within normal and handicapped populations of school children (Bower, 1960; Chamberlain, 1966; Miller, 1972; Ross, Lacey, and Parton, 1965; Spivak and Swift, 1966; Walker, 1970; Werry and Quay, 1971; Zax, Cowen, Izzo, and Trost, 1964). Diagnostically these instruments have proven valuable in the specification of relevant areas of behavioral intervention.

Past research in the study of the relationship between pupil classroom behavior and reading achievement has utilized two primary means of assessing student behavior, (1) teacher ratings, and (2) direct classroom observation. A critical question then that must be addressed is,

what is the relationship between teacher judgments of pupil behavior and such techniques as observation by external observers?

Young-Maston (1938) appears to be the first investigator to study the accuracy of teacher judgments of child behavior through the use of observational data. Observational data derived from continuous diary recording of classroom behavior did not discriminate between pupils identified by teachers as the worst behaved and averaged behaved, although teacher ratings of the students behavior produced highly significant differences between the two groups. In this study then, the accuracy of teacher judgment was not corroborated by classroom observation. Bolstad (1974), however, suggested that this study may have been confounded by the recording procedures used in analyzing student behavior by the outside observers.

Bolstad (1974) further investigated the relationship between teacher ratings and direct behavioral observations of children selected by their teachers as best behaved, average behaved, and least well behaved. Third and fourth grade teachers were asked to rate these pupils on three retrospective rating scales measuring classroom conduct. A behavioral coding system (Walker, Mattson, and Buckley, 1971) which consisted of four appropriate and four inappropriate behavior categories was used by professionally trained observers. Results of the study indicated that teachers perceived highly significant differences between the three groups on all three rating scales. Behavioral observation data also discriminated among the three groups of pupils. The proportions of appropriate classroom behavior were .92, .88, and .70 for the best, average and least well behaved groups, respectively. Boys averaged .80 and girls .87 across the three groups, a difference that was also significant. These results strongly suggested that teachers in his sample were quite

accurate in their assessment of pupil behavior in their classes.

Nelson (1971) utilized direct observational procedures to investigate differences between children initially classified as conduct-disturbed or normal based upon behavioral ratings provided by their regular education teachers. The specific purpose of this study was to confirm the usefulness of teacher ratings and direct observation by validating one technique against the other. The results indicated that the normal subjects engaged in significantly more task-oriented behavior and significantly fewer deviant behaviors than the previously identified conduct-disturbed students, and that results of the observation procedures corroborated the accuracy of the teachers' initial ratings.

Jones and Cobb (1973) conducted a study in which they compared the relative usefulness of teachers' versus observers assessments of pupil behavior. Specifically the objectives of their study were to measure the convergent and discriminant validity of observational data versus teacher ratings and to evaluate the predictive validity of these two data sources against standardized achievement scores. Teachers rated eighteen categories of classroom behavior on a seven point scale; observers rated the same categories using an interval recording procedure. The Stanford Achievement Test was administered to all pupils to measure reading achievement. The results indicated that teacher ratings of students' classroom behavior tended to be more predictive of reading achievement than the observations of the same behavioral categories, e.g. validity coefficients of .62 versus .39. As might be expected, however, the convergence between teacher ratings and behavioral observation data was weak, with coefficients of .28 and .27 for appropriate and inappropriate behaviors respectively. Their results do suggest that when teachers are

given similar criteria for judging behavior, as are professionally trained observers, they are at least as skilled as observers in their judgments of child behavior, and more skilled in judging classroom behaviors predictive of achievement.

Camp and Zimet (1974) examined the relationship between two measures of teacher ratings of student behavior, the Pittsburgh Adjustment Scale (Ross et al., 1965) and the Connors Teacher Rating Scale (Conners, 1969) with independent observations of classroom behavior and achievement test scores in two first grade classrooms. Correlations were performed between classroom observation variables and those subscales of the PASS and CTRS which did not show any significant differences between the two classrooms. Observation of off-task behavior correlated significantly and positively with the Daydreaming subscale of the CTRS, .31 and significantly by negatively with the Presocial subscale of the PASS, -.38. Relationships were further examined between teacher ratings and observations within each classroom. The findings indicated that the two teachers varied greatly in the extent to which their ratings correlated with the ratings of the observers. Camp and Zimet reported that one teacher had nine out of twenty-nine significant correlations, while the other teacher had nineteen out of twenty-nine. Unfortunately, none of the correlations were reported. With the small number of teachers sampled, and the high degree of variability between the two, the findings appear to be inconclusive regarding the relationship between teacher ratings and observational data. In relation to achievement test scores, the authors gave no indication as to which method of behavior recording, teacher judgment or direct observation, was most closely related to measured academic achievement.

Several reasons may account for the difference noted between behavioral assessment by means of teacher ratings and direct observation. It appears that one significant difficulty is that the items on the teacher rating scale may not correspond to the items on the instrument used by the observer. Even when the items are comparable, the situations in which the child is observed may not. Thus, even if observations are made by well-trained, unbiased persons, the observations inevitably reflect a significantly smaller segment of behavior than that encountered by the teacher. Furthermore, both teacher and observer ratings may vary in the extent to which they report positive as opposed to negative behavior. Werry and Quay (1969) suggest that most children, whether they have behavior problems or not, are on-task seventy-five percent of the time. Bolstad's (1974) findings appear to concur with this hypothesis. Camp and Zimet (1974) suggest that because most children are on-task for the majority of the time, positive behavior will be more likely to be observed by both external observer and teacher since it is more common. Conversely, if a child's behavior is truly negative or deviant and leads to explicitly stated rules over twenty-five percent of the time, it should also be readily observed by both teacher and observer.

CHAPTER III

DESIGN OF THE STUDY

Population and Sample

Subjects for the present study were 93 students from a metropolitan elementary school in mid-Michigan. The school was selected on the basis of its representative population in terms of race and socioeconomic status according to school district administrators. Support for the study was solicited from the school principal and ten teachers, nine female and one male, who voluntarily indicated a willingness to participate in a larger, related research project (Blom, et al., 1980).

The school is located in a "blue collar" area of the city and district sources indicated that approximately 21% of the students' families are currently receiving aid to dependent children support and that this figure is slightly lower than the district average. Twenty-six percent of the students are of minority racial status, as compared to a district wide average of 35%. Seventy-four percent of the school population is identified as Caucasian, 17% Black, 6% Latino, and 3% Asian. Specific socioeconomic and racial data on individual children were unavailable.

Regarding the educational status of the students at the school, school district sources indicated that 27.5% of the students are receiving Title I reading and mathematics support services as compared to

a 29% district-wide average. Four percent of the school's students are receiving special education services as compared to an average of 5.4% district-wide.

Table I indicates the average percentile reading score on the Stanford Achievement Test (SAT) for the school's students as compared to district-wide elementary totals.

TABLE I

Average Percentile Rank - SAT
Reading for Grades 1-6

	1977-78	1978-79	1979-80	1980-81
The School	45.9	49.5	48.6	49.1
District	49.6	51.3	51.9	53.4

All students in two classrooms of each grade level, two through six, were initially evaluated. That population consisted of 226 students, 124 males and 102 females. Male-female totals at each grade level ranged between 34 and 51, and it was possible to sample at least 16 boys and 15 girls at each grade level. Subjects for whom not all relevant data (behavior checklist and intelligence) was available, or students whose protocols were reported to be invalid by test examiners or test scorers, were excluded from further involvement in the study.

Ninety-three students from the original group were subsequently identified for longitudinal follow-up three years after the initial evaluations. These subjects were those (1) who were able to be identified as

still attending the school district after the three year interval, (2) whose parents or guardians consented to release of their child's reading achievement test scores from the school district records, and (3) whose SAT scores were available for all four years.

The mean age and grade levels of the final sample were 114.8 months ± 16.2 months, and 4.0 ± 1.5 respectfully. Fifty-three percent of the sample were male and 47% female. The mean intelligence quotient for the sample was 104.3 ± 11.8. A breakdown by grade in Table 2 suggests a rather equal representation of students at each level.

TABLE 2

Relative Percentage and Absolute Frequency of Subjects at Each Grade Level

Grade	Frequency	Relative Percentage
2	17	18.3
3	19	20.4
4	21	22.6
5	17	18.3
6	19	20.4

A review of relative demographic, cognitive, and behavioral data suggested that the 93 subjects appears to be highly representative of the original population of students evaluated. The sample was comprised of 53% (n=49) males and 47% (n=44) as compared to 54.9% males (n=124) and 45.1% (n=102) in the original population. Table 3 indicates a relative comparison of the 93 students to the original population in terms of

average age, grade level, IQ, and raw scores on the behavior dimensions identified for analysis in the present study.

TABLE 3

A Comparison of Relevant Sample Versus School Population Characteristics

	Sample		Population	
	Mean	S.D.	Mean	S.D.
Age (months)	114.8	16.2	116.4	17.6
Grade Level	4.0	1.5	4.2	1.4
IQ	104.3	11.8	101.3	12.7
Behavior Scales				
Attention/Motivation	16.5	4.3	16.8	4.4
Disrespect/Defiance	9.5	2.2	12.4	2.8
Attention-Seeking	10.6	2.4	10.4	2.2
Academic Self-Confidence	6.8	1.2	7.1	1.4
Socialization	3.5	0.9	3.6	1.0
Externalized Aggression	9.5	2.2	9.5	2.3

Appendix A indicates a breakdown of mean Stanford Achievement Test Scores by a grade and sex of student, while Appendix B indicates a breakdown of mean behavior scale raw scores also by grade and sex of student.

Instrumentation

Dependent variables assessed in this study were teacher ratings of student classroom behavior, sex of student, and intelligence. Reading achievement was the dependent variable.

Student Behavior. The School Behavior Checklist: Form A2 (SBCL) (Miller, 1972) was used to assess student classroom behavior. The SBCL is an inventory of behaviors designed to help teachers communicate their impression of children in their classrooms. This inventory covers a wide range of social and academic behaviors from moderate social deviance indictative of pathology to social and academic competence. Rare and extreme deviant behaviors were not included in the inventory. The SBCL provides a relatively simple procedure for acquiring teacher's impressions of their students. The SBCL is such that achievement-related versus more mental health-related behaviors could be identified.

The SBCL was developed by Miller (1972) primarily from cross-validation of an earlier measure, the Pittsburgh Adjustment Survey Scales (Ross, Lacy, and Parton, 1965) with the addition of fourteen learning disability items, and two manifest anxiety items. The standardization sample was comprised of 2,627 male and 2,746 female students in the city, county, parochial and private schools of Louisville and Jefferson County, Kentucky. Teachers who rated the protocols had on the average 13.8 years of experience.

Form A2 consists of 96 statements regarding childrens' behavior. Teachers were asked in a forced choice format to decide if the behavior described the child being rated, and if so, check True (T), if not, False (F). Six factor scales and one clinical scale have been derived for the SBCL-A2. The factors were found to be identical for boys and girls.

Test-retest reliability coefficients for the individual subscales range from .70 to .90 with a reliability coefficient of .93 for the Total Disability Scale. Test-retest reliability is similar with the individual subscales ranging from .70 to .89 with an overall test-retest coefficient of .89.

The student behavior ratings used as independent variables in the prediction of future reading achievement in this study consisted of new behavioral dimensions derived from SBCL items which were found to be significantly related to reading achievement for the total sample and boys and girls independently. Appendices D through F indicate all attained simple correlations.

The basic rationale behind creation of these behavior dimensions or scales was to reduce the number of behavior variables to be considered as independent variables in the prediction of later reading achievement. Scales derived from the original SBCL sample were deemed not meaningful for purposes of the present study as they appeared too heterogenous in nature, and did not reflect specifically those behavioral dimensions which have been found to be related to reading achievement. Quay (1965) has suggested that the behavior of children may be most profitably viewed as dimensions should satisfy three internal criteria: (1) that they are mutually exclusive as possible, (2) that they are objective as possible, and (3) that they are specifically demonstrated as a cluster.

Kerlinger and Pedhazur (1973) and Kim and Kohout (1975) state that data-reduction procedures, such as scale development, are practically and statistically important in problems involving prediction in general, and multiple regression procedures in specific. Both sets of authors suggest that the creation of new variables, which are composite scales of highly intercorrelated items, is one solution to the problem of multicollinearity

(Gordon, 1968). Multicollinearity refers to the situation in which some or all of the independent variables in a prediction equation are highly intercorrelated. Prediction, and subsequent explanation of the predictive relationship between the independent variables and dependent variable is enhanced when the predictor variables of interest are essentially independent or orthogonal (Kerlinger and Pedhazur, 1973).

Forty-nine items from the SBCL, that were found to be significantly related to initial levels of reading achievement, formed the pool of items in the development of the new behavior scales which were to be used in the prediction of future achievement. This item pool was subsequently subjected to factor analytic procedures to examine the underlying statistical interrelationships of the items and to reduce the number of behavioral predictor variables to a manageable number of statistically and conceptually meaningful behavioral dimensions.

Four to eight factor solutions were examined through principal factoring with iteration and varimax rotation procedures and factors were further analyzed for internal consistency reliability. Extraction of six factors produced what appeared to be the most parsimonious, independent, and conceptually meaningful behavioral scales given the previously identified item pool.

For the purpose of identifying items in the formation of the scales, the following guidelines were observed:

1. No item was to be included in a scale unless it had a factor loading greater than or equal to .30 with a loading of .40 preferable.
2. No item was to be included on more than one scale.

3. Each scale should have a minimum of three items (a minimal requirement in terms of measuring internal consistency reliability).

4. The items included in each scale should be conceptually related.

The first step in the development of the scales was to examine the factor loadings of each item for the four to eight factor solutions. All factor loadings of .30 or greater were identified and each item was examined to determine which one factor the item was best represented by. Items were selected for a given factors based on the highest loading value of that item across factors. All derived factors were subsequently evaluated in terms of their internal consistency reliability.

The next step in the procedure was to examine the factors in terms of their conceptual clarity. As was indicated, the six factor solution clustered the items in what appeared to be the most conceptually meaningful as well as statistically parsimonious manner. All six factors contained three or more items and all were comprised of items which appeared theoretically related. Appendix C indicates the item loading values for the six factor solution.

The final step in the process was to examine the items within each scale and devise a descriptive title for each. The following is a conceptual and statistical description of the six derived scales.

SCALE I - ATTENTION/MOTIVATION

Scale I is comprised of thirteen items which tend to describe a student's general adaption to school structure and expectations (e.g. works well alone, is interested in school work, finishes classroom assignments, does homework, is sure of self), as well as a student's ability to attend to task (e.g. is alert in class, is able to conceptrate on things, is distractable). The internal consistency reliability coefficient

of this scale was .94.

SCALE II - DISRESPECT/DEFIANCE

Scale II consists of ten items which tend to described a student's active non-compliance to authority and structure (e.g. will put up an argument when told not to do something, is infuriated by any form of discipline, is stubborn, sulks when things go wrong), as well as externalized emotional lability (e.g. is easily upset by changes, has changeable moods, when angry will do things like slamming the door or banging the desk). The reliability coefficient of this scale was .87.

SCALE III - EXTERNALIZED AGGRESSION

Scale III is comprised of eight items which reflect a student's verbal and physically aggressive interpersonal orientation with peers in the school setting (e.g. starts fighting over nothing, hits and pushes other children, teases other children, uses abusive language towards other children, threatens to hurt other children when angry). The reliability coefficient of this scale was .87.

SCALE IV - ATTENTION-SEEKING

Scale IV is comprised of nine items which describe a student's "immature", impulsive, active, attention-seeking behaviors (e.g. likes an audience all the time, never seems to be still for a moment, interrupts whoever is speaking, disturbs other children with boisterous behavior, does things which are normal for children much younger, does things just to attract attention). The reliability coefficient of this scale was .87.

SCALE V - ACADEMIC SELF-CONFIDENCE

Scale V consists of six items which tend to describe a student's more passive, reticent orientation towards academic and classroom demands (e.g.

approaches a task with an air of defeatism, does not volunteer to recite in class, seems dull and slow to catch on, will not ask questions even when unsure of how to do the work, has difficulty speaking when excited or upset). The internal consistency coefficient of this scale was .60.

SCALE VI - SOCIALIZATION

Scale VI is comprised of three items which reflect a student's social integration with his peers (e.g. is popular with classmates, prefers to play alone and be alone, is slow in making friends). The internal consistency coefficient of this scale was .73.

Table 4 indicates the six new behavioral scales with their corresponding items and factor loadings in descending magnitude for each scale.

TABLE 4

Reading Achievement-Related Behavioral Scales

Item No.	Item Description	Factor Loading
	SCALE I: ATTENTION/MOTIVATION	
22	Does homework	.88
18	Fails to carry out tasks	.84
19	Lacks ambition to do well in school	.80
76	Finishes classroom assignments	.80
36	Finds it hard to study	.79
52	"Drags feet" when requested to do something	.72
57	Distractible; can't concentrate	.68
64	Is interested in school work	.65
48	Is able to concentrate on things	.61
7	Is alert in class	.57
2	Tends to give up if has something hard to finish	.53
27	Is sure of self	.50
38	Works well alone	.48

"TABLE 4 (cont'd)."

	SCALE 2: DISRESPECT/DEFIANCE	
54	Sulks when things go wrong	.68
47	Argues with the teacher	.64
21	Will put up an argument when told not to do something	.62
29	Has changeable moods	.59
77	Gives other children dirty looks	.51
26	Is easily upset by changes	.51
39	When angry, will refuse to speak to anyone	.47
62	Is stubborn	.47
86	When angry, will do things like slamming the door or banging the desk	.43
34	Is infuriated by any form of discipline	.38
	SCALE 3: EXTERNALIZED AGGRESSION	
13	Hits and pushes other children	.79
20	Does things to get others angry	.64
23	Teases other children	.64
31	Does not respect other people's belongings	.56
17	Is not considerate of others	.55
72	Threatens to hurt other children when angry	.51
5	Starts fighting over nothing	.48
28	Uses abusive language towards other children	.48
	SCALE 4: ATTENTION-SEEKING	
35	Likes an audience all the time	.77
51	Tries to be the center of attention	.72
70	Does things which are normal for children much younger	.58
87	Acts in a fearless, "dare-devil" manner	.53
92	Disturbs other children with boisterous behavior	.52
46	Never seems to be still for a moment	.47
3	Interrupts whoever is speaking	.45
66	Does things just to attract attention	.45
78	Deliberately interrupts what is going on by asking silly questions	.43
	SCALE 5: ACADEMIC SELF-CONFIDENCE	
43	Seems dull; slow to catch on	.51
44	Will not ask questions even when unsure of how to do the work	.47
16	Approaches a task with an air of defeatism	.46
63	Never speaks up even when there is cause to be angry	.44

"TABLE 4 (cont'd)."

88	Has difficulty speaking when excited or upset	.44
12	Volunteers to recite in class	.37
	SCALE 6: SOCIALIZATION	
79	Is slow in making friends	.85
69	Is popular with classmates	.71
75	Prefers to be alone and play alone	.33

Table 5 indicates the intercorrelations among the six identified behavior scales. As can be seen, relatively high positive correlations are evidenced among the Disrespect/Defiance, Externalized Aggression, and Attention-Seeking Scales. One explanation of the high degree of relationship among these scales would appear to be that all items within the scales share common variance with reading achievement and were selected on that basis. Further, numerous scale items loaded highly on more than the single factor for which they were identified (e.g. Is stubborn, argues with teacher, does things to attract attention, and threatens other children). Conceptually, such items would appear appropriate to each factor on which they loaded.

The Academic Self-Confidence and Socialization Scales appear relatively independent of the other four scales. The relatively high intercorrelations of the first four scales would suggest that their contribution to the prediction of future reading achievement will be diminished relative to the degree of multicollinearity among the scales.

TABLE 5

Intercorrelation of the Behavior Scales

Scale	1	2	3	4	5	6
Attention/ Motivation	-					
Disrespect/ Defiance	.42	-				
Externalized Aggression	.38	.64	-			
Attention-Seeking	.44	.66	.66	-		
Academic Self-Confidence	.34	.08	.09	.08	-	
Socialization	.47	.15	.15	.17	.23	-

Reading Achievement

Reading achievement was assessed by means of the reading subtests of the Stanford Achievement Test (SAT) (Madden et al., 1975). Appropriate forms of the test were administered: Primary Level II for middle second to beginning third graders, Primary Level III for third to beginning fourth graders, Intermediate Level I for fourth to beginning fifth graders, Intermediate Level II for fifth to sixth graders, and Advanced for seventh through ninth graders.

Standardized total reading scores were used for purposes of analysis in the present investigation. Students' total scores are composed of various subscales depending upon the level of the test. Only Reading Comprehension and Word Study Skills are used to make up the Total Reading score through the Intermediate II level. The Vocabulary score is not included

at these lower levels as this particular subtest is dictated by the teacher and not read by the student. At the Advanced level, Vocabulary is read by the students and Word Study Skills is not included; hence, the Total Reading score of the Advanced level includes Vocabulary and Reading Comprehension subtests.

Standardized scaled scores on the SAT were obtained through a computerized application of Thurstone's absolute scaling procedure, "resulting in a system of interbattery standard scores, which permits the translation of raw scores at each level into scaled scores with comparability across levels for a test area," (Madden et al., 1975, p. 30). The authors point out that within a single test area, such as Reading, the scaled scores obtained by the aforementioned methods are thus directly comparable from grade to grade, battery to battery, and form to form.

The SAT was selected as the primary criterion variable for reading achievement in the present study as individual students' annual scores were readily attainable from the Lansing School District, as well as the fact that the SAT is a widely-used measurement tool in school districts throughout the country. Evaluation of reported technical data indicates reasonable measurement characteristics as well as ample sampling and norming procedures. Internal consistency or split-half reliability is reported at .86-.96 across all forms and levels for the reading subtests.

Intelligence. For purposes of the present study, the Otis-Lennon Mental Ability Test (Otis and Lennon, 1969) was employed as a measure of general aptitude or intelligence. This test is a widely used, group administered general mental ability measure which evaluates students' facility in dealing with abstract concepts presented in verbal, figural, or symbolic form. Emphasis is placed upon the ability to discern

relationships of varying complexity within these three symbol systems.

Elementary Level I was used in the second and third grades and Elementary Level II was used in grades four through six. Both forms have reasonable proven measurement characteristics. Alternate-form reliability coefficients are reported as ranging from .85 to .93 and internal consistency reliability coefficients ranged from .88 to .95. Scores are reported as a deviation IQ. Deviation IQ's are normalized standard scores with a mean of 100 and a standard deviation of 16. A measure of general aptitude or intelligence was included as an independent variable in the present study, as IQ has been shown to contribute to a significant amount of the variance in elementary reading achievement scores as well as reported student behavior (Miller, 1972).

Definitions

In order to determine the relative accuracy by which the selected independent variables predict reading achievement, students were classified Good, Average, or Poor Readers based upon their attained standard score on the Stanford Achievement Test. A Poor Reader was defined as a student who scored less than or equal to one standard deviation below the mean on the SAT according to test norms (T-Score 40). An Average Reader was defined as a student who scored within plus and minus one standard deviation above the mean on the SAT (T-Score 60). The aforementioned cut-off criteria were selected as they (at the lower achievement levels) closely approximate the school district's decision rule for considering a student for compensatory or remedial reading support services.

An important aspect of the study was to establish the relative predictive accuracy of the independent variables with students at varying achievement levels. If long range educational decisions are to be made

based upon the prediction batteries, it would appear imperative that errors in prediction be negligible, especially for students in the extreme groups, Good and Poor Readers, for whom special educational programs might most likely be instigated.

Data Collection

Data for the initial sample of second through sixth grade students were collected during the course of a larger project of which this researcher was intimately involved (Blom, et al., 1980). This initial research project was funded through an all-university research grant at Michigan State University and supported through the auspices of Departments of Psychiatry and Special Education.

All instruments were administered within the first quarter of the academic school year. School had been in session approximately two months, giving the teachers adequate time to become acquainted with their students, an important aspect in terms of the reliability and validity of their judgments regarding pupil behavior on the SBCL.

The Otis-Lennon Mental Ability Test was administered over the course of a two week period by a group of undergraduate special education students, who were initially solicited, briefed and trained by this researcher in the research test administration procedures. Participating teachers were given two weeks to complete the School Behavior Checklist on each of their students. Protocols were hand-scored and recorded by this researcher and the undergraduate volunteers.

Reading achievement data was obtained through the Evaluation Services Department of the participatory school district. The Stanford Achievement Test is administered annually in the district, and has in the past been used as a primary screening device for identification of students, second

through ninth grade, who may be in need of supplemental compensatory support in reading and mathematics. Reading achievement data for each of the three subsequent years after the initial study was obtained on each of those students who continued to reside in the Lansing School District, and for whom parental release of information forms could be obtained.

Statistical Hypotheses

There were five major hypotheses formulated and tested in the present study of the relationship between student behavioral and cognitive characteristics and reading achievement in the elementary grades. The first hypothesis was concerned with identifying those student behavioral dimensions which were related to reading achievement over time. The second hypothesis evaluated the relative contribution of the behavioral dimensions to the explanation of future reading achievement while taking into account I!, sex, and initial level of achievement. Finally, the third through fifth hypotheses were concerned with evaluating the relative accuracy and utility of the identified behavior scales independent of and in conjunction with IQ, sex, and initial level of reading achievement (SAT_1) in predicting future reading achievement.

Ho_1: There is no linear relationship between the identified behavior scales and reading achievement as measured by the Stanford Reading Achievement Test:

a. After one year (SAT_2).

b. After two years (SAT_3).

c. After three years (SAT_4).

Ho_2: The identified behavior scales have no linear effect on SAT_2, SAT_3, and SAT_4 adjusting for the effects of IQ, sex of student, and initial level of reading achievement (SAT_1).

Ho_3: There is no linear relationship between initial level of reading achievement, IQ, sex of student, and the identified behavior scales and reading achievement as measured by the Stanford Reading Achievement Test:

a. After one year (SAT_2).

b. After two years (SAT_3).

c. After three years (SAT_4).

Ho_4: The linear rule derived from the identified behavior scales does not predict the achievement scores of good, average, and poor readers:

a. After one year (SAT_2).

b. After two years (SAT_3).

c. After three years (SAT_4).

Ho5: The linear rule derived from initial level of reading achievement, IQ, sex of student, and the identified behavior scales does not predict the achievement of good, average and poor readers:

a. After one year (SAT_2).

b. After two years (SAT_3).

c. After three years (SAT_4).

Data Analysis Procedures

All data in the study were analyzed at the Michigan State University Computer Center using various programs from the Statistical Package for the Social Sciences (SPSS), Nie, et al. (1975).

Items from the SBCL that were and were not related to reading achievement were identified by means of generating a series of Pearson Product-Moment zero-order correlation coefficients. An alpha level of .01 was set as the decision criterion.

Those items from the SBCL which were found to be significantly related to reading achievement (SAT_1) for the total sample or those items which were significantly related to reading achievement (SAT_1) for boys or girls only, were subsequently subjected to further analysis procedures. These items formed the pool of items for the development of behavioral scales which were to be used to predict future reading achievement (SAT_2 to SAT_4). Factor analytic statistical procedures were utilized to reduce the number of individual behavioral predictor variables to a manageable number of statistically and theoretically meaningful behavioral dimensions such as to improve the precision of prediction of future achievement.

Initial principal factoring with iteration allowed, (1) all factors imposed to be orthogonal or independent, an important consideration in terms of future prediction equations, (2) factors to be arranged in the order of their importance, and (3) the first factor to be a general factor, that is, having significant loading on every variable. This factoring method automatically replaced the main diagonal elements of the correlation matrix of all the variables with communality estimates. Initial estimates of the communalities are given by the squared multiple correlation between a given variable and the rest of the variables and subsequently tended to improve the theoretically meaningful patterning of the variables (Nie, et al., 1975). For purposes of the present study, extraction of six factors produced the most parsimonious, independent, and conceptually meaningful behavioral dimensions or scales, given the identified item pool.

These six identified behavior dimensions were further subjected to item and scale analysis procedures by means of the Reliability subprogram of SPSS. This program provides a means for evaluating multiple-item

scales through the computation of widely recognized coefficients of reliability. For purposes of the present analysis, Cronbach's Alpha Coefficient was computed for each of the identified behavioral scales, and item-scale correlations were evaluated to determine the item's relative contribution to the variance of the scale. Internal consistency (coefficient alpha) reliability indicates the degree of measurement error due to measurement inconsistencies of the scale items. The program allows one to observe changes in the overall reliability of a scale with individual items excluded as well as included in that scale.

Partial correlation coefficients were computed in order to ascertain the relative contribution of the behavior scale to future reading scores while adjusting for the effects of IQ, initial reading achievement, and sex of student, which are all postulated to all be highly and significantly related to future reading achievement. The Partial Correlation subprogram of SPSS indicates the zero-order correlation, the degree of freedom, and a measure of the statistical significance of the relationship. A .01 alpha level was set as the decision criterion.

Multiple regression analysis was the procedure used to determine the accuracy by which future reading achievement was predicted by the behavior scales, independent from and in conjunction with IQ, initial level of reading achievement, and sex of student. Multiple regression is a general statistical technique through which one can analyze the relationship between a dependent or criterion variable and a set of predictor or independent variables. As a descriptive tool, the primary purpose of multiple regression analysis is, (1) to find the best linear prediction equation and evaluate its predictive accuracy, (2) to control for other confounding factors in order to evaluate the contribution of a specific

variable or set of variables; and (3) to find structural relations and provide explanations for seemingly complex and multivariate relationships (Kim and Kohout, 1975, p. 321). Kerlinger and Pedhazur (1973) state that multiple regression is most often the best method of analysis of non-experimental data as it is "closely related to the basic purpose of science, the explanation of natural phenomena." A significant strength of multiple regression for purposes of the present study was its ability to handle categorical as well as continuous independent variables. Further, predicted reading achievement scores could be generated and compared against actual scores to help determine not only the predictive accuracy, but predictive utility of the derived linear model, or prediction equation.

Stepwise multiple regression procedures were employed to determine those variables that would yield an optimal equation for the purposes of accurately predicting future reading achievement scores. Two separate equations were generated for predicting future achievement (SAT_2, SAT_3 SAT_4). In the first equation only the behavior scales were entered, while in the second equation the behavior scales were entered with IQ, initial level of reading achievement ($RACH_1$), and sex of student. In utilizing the stepwise procedure, the independent variables were entered into the prediction equation such that the variable that explained the greatest amount of variance in the dependent variable entered first; the variable that explained the greatest amount of variance in conjunction with the first entered second and so on. Regression coefficients or beta weights for each of the independent variables were generated and evaluated statistically in terms of an F ratio. At each step of the analysis, F ratios were computed for variables not yet in the equation. For purposes of this study, all variables were allowed to enter into the equations in order to maximize the predictive power of the equation.

In order to further assess the ability of the independent variables of interest in this study to accurately predict future reading achievement, predicted scores (Y'), and actual scores (Y) and residuals (i.e. errors in prediction) were calculated for each student. A decision rule was then evoked to determine whether the predicted score was commensurate with the actual score. For purposes of this study, a predicted score within plus or minus one standard error of measurement of the actual score, was considered a correct diagnostic classification or "hit", while a predicted score outside of plus or minus one standard error of measurement was considered a diagnostic "miss". These diagnostic "hits and misses" were subsequently incorporated into three by three detection tables as per Feshbach, et al., (1974) and Satz and Friel (1974). Six separate detection tables were generated, one for each prediction equation at each measurement of future reading achievement (SAT_2, SAT_3, SAT_4). The following illustration defined the nature of the detection table (the actual tables included good, average, and poor readers):

Actual Outcome

Predicted	Poor Reader	Good Reader	Total
Poor Reader (+)	Valid Positive Rate	False Positive Rate	
Good Reader (-)	False Negative Rate	Valid Negative Rate	
Total			

The pratical utility of the prediction equations were finally computed directly by the principle of inverse probability, or Bayes Theorem (Satz and Fletcher, 1979). For each detection table, base rate information was given as the total number of actual poor, average, and good readers. Table 6 indicates the base rates for the present sample.

TABLE 6

Relative Base Rates of Poor, Average, and Good Readers for SAT_2, SAT_3, and SAT_4

	SAT_2	SAT_3	SAT_4
Poor Readers	14.0%	12.9%	11.8%
Average Readers	71.0%	73.1%	75.3%
Good Readers	15.0%	14.0%	12.9%

Based of this base rate information, conditional probability values for each of the predictive test signs were determined. For example, the probability (P) that an individual student would become a poor reader (PR), given a positive or poor sign (+) was calculated by the following Bayesian expression:

$$P(PR +) = \frac{P(PR) \cdot P(+ PR)}{P(PR) \cdot P(+ PR) + P(GR) \cdot P(+ GR)}$$

The practical implications of relatively high errors in prediction (i.e. high false-positive and false-negative rates) are rather straightforward for both schools and individual students. If a preventive or early intervention program were established for students with high-risk test

signs (i.e. predicted poor achievement) and high false positive rates in prediction were evidenced, wherein many of the students turned out to be average to superior readers without any intervention, the sum result would represent unnecessary expense and effort to the school as well as potentially serious implications for the student who is incorrectly "labeled" and unnecessarily targeted into "lower track", remedial programs. On the other hand, the more potentially detrimental error in terms of long term implication for the individual student, the school, and community as a whole may be the false negative detection rates, wherein large numbers of students may be predicted to do well, when in fact the contrary is true. As such, a high proportion of high-risk students, who might strongly benefit from early intervention or remediation, would be systematically excluded from such services based on the prediction formula.

Past research and practical consideration would suggest that prediction error rates (false positives and false negatives) should be less than 10%, and the generated prediction equations in the present study were evaluated in terms of this criteria.

CHAPTER IV

RESULTS

To summarize, the study incorporated a descriptive, longitudinal design which purported (1) to identify teachers' perceptions of behavioral attributes of elementary age students which were highly and significantly related to reading achievement over time, and (2) to investigate the efficiency and utility of using teachers' ratings of student behavioral attributes to predict future reading performance. Teachers' ratings of student behavior were evaluated independently and concurrently with IQ, initial level of reading achievement and sex of student as to their ability to predict subsequent reading achievement after one, two, and three year intervals.

Results of the study are reported in relation to the stated five statistical hypotheses.

Hypothesis I: There is no linear relationship between the identified behavior scales and reading achievement after one, two, and three year intervals.

All of the behavior scales were found to be highly and significantly related to reading achievement at all time intervals (P .01). The Attention/Motivation Scale proved to be the most highly related to reading achievement of all of the scales at the initial as well as subsequent one year intervals (-.38 to -.49). Attained correlational ranges for the remaining five scales included Disrespect/Defiance (-.33 to -.37), Attention-Seeking (-.34 to -.36), Academic Self-Confidence (-.24 to -.36), Externalized

Aggression (-.30 to -.36), and Socialization (-.23 to -.39). Table 7 indicates the relative simple correlations between the behavior scales and reading achievement over four years.

TABLE 7

Correlations Between the Behavior Scales and Reading Achievement

Scale	SAT_1	SAT_2	SAT_3	SAT_4
Attention/ Motivation	-.49*	-.48*	-.40*	-.38*
Disrespect/ Defiance	-.33*	-.34*	-.35*	-.37*
Attention-Seeking	-.36*	-.34*	-.35*	-.35*
Academic Self-Confidence	-.36*	-.34*	-.24*	-.29*
Externalized Aggression	-.31*	-.32*	-.30*	-.36*
Socialization	-.33*	-.39*	-.29*	-.23*

$P < .01$

Hypothesis II: The identified behavior scales have no linear effect on SAT_2, SAT_3, and SAT_4 adjusting for the effects of IQ, sex of student, and initial level of reading achievement (SAT_1).

As is evident from Table 8, statistically controlling for sex of student evidenced little impact on the predictive relationship between the behavior scales and reading achievement. All correlations remained statistically significant ($P < .01$) with the exceptions of Academic Self-Confidence and SAT_3 and Socialization and SAT_4.

TABLE 8

Correlation Between Reading Achievement (SAT_2 to SAT_4) and the Behavior Scales Controlling for Sex of Student

Scale	SAT_2	SAT_3	SAT_4
Attention/ Motivation	-.46*	-.36*	-.36*
Disrespect/ Defiance	-.33*	-.34*	-.36*
Externalized Aggression	-.30*	-.27*	-.34*
Attention-Seeking	-.31*	-.31*	-.32
Academic Self-Confidence	-.33*	-.22*	-.28*
Socialization	-.37*	-.27*	-.21*

*$p < .01$

Because of measured student intelligence's (IQ) high correlations with reading achievement as well as its moderate to high observed relationship with the behavior scales, the attained initially strong correlations with reading achievement appeared appreciably diminished while controlling for the effects of IQ. Table 9 reflects those scales which maintained low, but significant ($p < .01$) relationships with reading achievement while taking in consideration their relative shared variance with IQ in explaining future reading achievement. Highest partial correlations were evidenced between Attention/Motivation and SAT_2, as well as Externalized Aggression and SAT_4.

TABLE 9

Correlation Between Reading Achievement (SAT_2 to SAT_4) and the Behavior Scales Controlling for IQ

Scale	SAT_2	SAT_3	SAT_4
Attention/ Motivation	-.29*	-.18	-.13
Disrespect/ Defiance	-.16	-.19	-.21
Externalized Aggression	-.24*	-.21	-.29*
Attention-Seeking	-.23	-.24*	-.23*
Academic Self-Confidence	-.17	-.03	-.09
Socialization	-.26*	-.14	-.03

*$p < .01$

Table 10 indicates the relatively weak partial correlation between the behavior scales and future reading achievement once the effects of initial level of achievement (SAT_1) had been considered. All partial correlations fell to nonsignificant levels.

TABLE 10

Correlation Between Reading Achievement (SAT_2 to SAT_4) and the Behavior Scales Controlling for SAT_1

Scale	SAT_2	SAT_3	SAT_4
Attention/ Motivation	-.12	-.01	-.02
Disrespect/ Defiance	-.11	-.15	-.20
Externalized Aggression	-.11	-.10	-.20
Attention-Seeking	-.05	-.10	-.12
Academic Self-Confidence	-.06	-.09	-.04
Socialization	-.22	-.05	-.03

Table 11 indicates that when the correlations between the behavior scales and reading achievement were controlled for by the relative effects of all three moderating variables, e.g. initial level of achievement, IQ, and sex of student, again as was the case with SAT_1 only, generally low, nonsignificant levels of relationships were evidenced.

TABLE 11

Correlation Between Reading Achievement (SAT_2 to SAT_4) and the Behavior Scales Controlling for SAT_1, IQ, and Sex of Student

Scale	SAT_2	SAT_3	SAT_4
Attention/ Motivation	-.11	.03	.02
Disrespect/ Defiance	-.09	-.14	-.17
Externalized Aggression	-.12	-.09	-.21
Attention-Seeking	-.05	-.09	-.12
Academic Self-Confidence	-.05	.11	.00
Socialization	-.21	-.03	.06

Hypothesis III: There is no linear relationship between initial level of reading achievement (SAT_1), IQ, sex of student, the identified behavior scales and reading achievement after one, two and three year intervals.

Forward stepwise multiple regression procedures were initially instigated using the behavior scales, only, as the independent predictor variables and SAT_2, SAT_3, and SAT_4 as the dependent measures. The behavior scales were entered in the regression solution 1) if they met minimal statistical criteria, and 2) in an order such that the scale that explained the greatest amount of variance in reading achievement was entered first, the scale that explained the greatest amount of variance in conjunction with the first, entered second, and so on.

Table 12 summarizes the relative contribution of the behavior scales to the variance in reading achievement after one year (SAT_2). In the aggregate, a multiple correlation (R) of .58 was attained, indicating that approximately 33% of the variance in reading achievement after one year could be explained by the contribution of the six behavior scales. The Attention/Motivation scale entered the equation first, itself contributing 23% to the total explained variance. Approximately 4% of the total variance was further explained by the Socialization scale and an additional 2% explained by the Attention-Seeking scale.

Table 13 recapitulates the relative contribution of the behavior scales to the variance in reading achievement after two years (SAT_3). Again, all six scales entered the equation, and a total multiple R of .49 was attained demonstrating that 24% of the total variance in reading achievement was attributable to teacher ratings of student behavior two years earlier. Again the Attention/Motivation scale proved to be the most predictive of the scales with a simple correlation of -.40, contributing 16% to the total explained variance in achievement. The Attention-Seeking scale explained an additional 4% of the total variance, while the remainder of the scales added negligible source of variance to SAT_3.

The relative contribution of the behavioral scales to the total explained variance in reading achievement after three years is summarized in Table 14. All six scales were entered into the prediction equation with a total multiple R of .51, explaining 26% of the variance in reading achievement after three years. The Attention/Motivation and Externalized Aggression entered at statistically significant levels ($p < .01$) and together accounted for 20% of the total explained variance in SAT_4. The Academic Self-Confidence scale loaded third and contributed an additional 3% to the total variance.

TABLE 12

Multiple Correlations of Reading Achievement After One Year (SAT_2) with the Behavior Scales

Variable	Simple R	Multiple R	R^2	R^2 Change	F to Enter	Significance	Beta	Constant
Attention/ Motivation	-.48	.4811	.2315	.2315	27.41	.00	-.44	85.29
Socialization	-.39	.5191	.2695	.0380	4.68	.03	-2.22	
Attention-Seeking	-.34	.5472	.2994	.0300	3.81	.05	-.36	
Academic Self-Confidence	-.34	.5661	.3205	.0210	2.72	.10	-1.28	
Externalized Aggression	-.32	.5720	.3271	.0067	.86	.36	-.42	
Disrespect/ Defiance	-.33	.5753	.3310	.0038	.49	.49	-.31	

TABLE 13

Multiple Correlations of Reading Achievement After Two Years (SAT_3) with the Behavior Scales

Variable	Simple R	Multiple R	R^2	R^2 Change	F to Enter	Significance	Beta	Constant
Attention/ Motivation	-.40	.3978	.1583	.1583	17.11	.00	-.33	79.25
Attention-Seeking	-.35	.4457	.1987	.0404	4.53	.04	-.48	
Socialization	-.29	.4688	.2198	.0212	2.41	.12	-1.59	
Disrespect/ Defiance	-.35	.4788	.2292	.0094	1.07	.30	-.46	
Academic Self-Confidence	-.24	.4866	.2368	.0076	.86	.36	-.72	
Externalized Aggression	-.30	.4880	.2382	.0014	.15	.70	-.22	

TABLE 14

Multiple Correlations of Reading Achievement After Three Years (SAT_4) with the Behavior Scales

Variable	Simple R	Multiple R	R^2	R^2 Change	F to Enter	Significance	Beta	Constant
Attention/ Motivation	-.38	.3775	.1425	.1425	15.13	.00	-.24	81.45
Externalized Aggression	-.36	.4468	.1996	.0571	6.42	.01	-.76	
Academic Self-Confidence	-.29	.4831	.2334	.0338	3.92	.05	-1.41	
Disrespect/ Defiance	-.37	.5029	.2530	.0195	2.30	.13	-.55	
Socialization	-.23	.5047	.2547	.0018	.20	.65	-.53	
Attention-Seeking	-.35	.5052	.2552	.0005	.06	.81	-.14	

Intelligence (IQ), sex of student, and initial level of reading achievement (SAT_1) were subsequently entered into the prediction equation with the behavior scales. As is quite evident in the regression summary in Table 15, inclusion of SAT_1 subsumed nearly the total explained variance in SAT_2, 78% of a total of 79%, with the socialization scale contributing an additional 1%, and the remaining behavior scales, IQ, and sex of student contributing no further to the explained total variance of SAT_2.

A separate prediction equation was computed to gain some insight into the relative contribution the explained variance in SAT_2 of the moderating variables SAT_1, IQ, and sex of student after the behavioral scales had been considered. Results again indicated that a multiple R of .58 was attained with the six behavior scales (33% of the total explained variance). Inclusion of IQ into the equation resulted in a multiple R of .73 (an additional 21% to the explained variance). Subsequent inclusion of sex of student added negligible explanation, while final inclusion of initial reading achievement (SAT_1) added an additional 25% to the explained variance in SAT_2 for a final multiple R of .89.

Table 16 summarizes the relative contribution of the behavior scales, IQ, sex of student, and SAT_1 to the total explained variance of reading achievement after a two year interval (SAT_3). SAT_1 again entered the prediction equation first with a simple R of .80, explaining 64% of the variance. All behavior scales, IQ, and sex contributed only 2% to the total explained variance in SAT_3 after inclusion of SAT_1. A final multiple R of .81 was attained, indicating that 66% of the total variance of SAT_3 was explained by SAT_1, the behavior scales, sex of student, and IQ.

TABLE 15

Multiple Correlations of Reading Achievement After One Year (SAT_2) With SAT_1, IQ, Sex of Student and the Behavior Scales

Variable	Simple R	Multiple R	R^2	R^2 Change	F to Enter	Significance	Beta	Constant
SAT_1	.88	.8833	.7802	.7802	322.94	.00	.81	14.30
Socialization	-.39	.8890	.7904	.0102	4.39	.04	-1.05	
Disrespect/ Defiance	-.34	.8902	.7925	.0021	.89	.35	-.12	
IQ	.68	.8908	.7935	.0010	.43	.52	.30	
Externalized Aggression	-.32	.8912	.7942	.0007	.29	.59	-.21	
Attention-Seeking	-.34	.8914	.7945	.0004	.15	.70	.10	
Sex	.18	.8915	.7947	.0002	.09	.76	-.32	
Academic Self-Confidence	-.34	.8916	.7949	.0002	.06	.81	-.81	
Attention/ Motivation	-.48	.8916	.7949	.0000	.02	.90	-.19	

TABLE 16

Multiple Correlations of Reading Achievement After Two Years (SAT_3) With SAT_1, IQ, Sex of Student and the Behavior Scales

Variable	Simple R	Multiple R	R^2	R^2 Change	F to Enter	Significance	Beta	Constant
SAT_1	.80	.8016	.6426	.6426	163.62	.00	.71	7.65
Disrespect/ Defiance	-.35	.8069	.6510	.0084	2.18	.14	-.32	
IQ	.63	.8088	.6541	.0031	.79	.38	.69	
Academic Self-Confidence	-.24	.8106	.6571	.0030	.77	.38	.39	
Sex	.20	.8119	.6592	.0021	.54	.47	.98	
Attention/ Motivation	-.40	.8124	.6600	.0008	.20	.66	.11	
Socialization	-.29	.8133	.6614	.0014	.35	.56	-.47	

Again a separate prediction equation was computed to observe the relative contribution of the moderating variables SAT_1, IQ, and sex of student to the explained variance of SAT_3 after the behavior scales had been considered. Twenty-four percent of the explained variance in SAT_3 was attributable to the behavior scales (multiple R = .49) with Attention/Motivation entering the equation first with a simple R of .40 (16% of the total explained variance). Subsequent inclusion of sex of student added 1% to the total explained variance. Further inclusion of IQ increased the multiple R to .68 (46% of the total explained variance), and final inclusion of SAT_1 increased the multiple R to .81 (66% of the total explained variance in SAT_3).

A summary of the relative contributions of the behavior scales, IQ, sex of student, and SAT_1 to the total explained variance of reading achievement after three years (SAT_4) is presented in Table 17. Again, SAT_1 entered the prediction equation first with a simple R of .75, explaining 56% of the variance. IQ entered the equation second at a .01 significance level, contributing an additional 4% to the total explained variance of SAT_4. The six behavior scales subsequently entered the equation, and contributed only an aggregate 1% to the explained variance after SAT_1 and IQ had been considered. A final multiple R of ,79 was attained indicating that 62% of the total variance of SAT_4 was explained by SAT_1, IQ, the behavior scales, and sex of student.

The relative contribution of the behavior scales to the total explained variance of SAT_4 prior to inclusion of the moderating variables was computed through a separate prediction equation. Twenty-six percent of the explained variance in SAT_4 was attributable to the behavior scales (multiple R = .51) with Attention/Motivation entering the equation first with a simple R of .38 (14% of the total explained variance). Externalized

TABLE 17

Multiple Correlations of Reading Achievement After Three Years (SAT_4) With SAT_1, IQ, Sex of Student and the Behavior Scales

Variable	Simple R	Multiple R	R^2	R^2 Change	F to Enter	Significance	Beta	Constant
SAT_1	.75	.7451	.5552	.5552	113.57	.00	.50	12.34
IQ	.68	.7702	.5931	.0380	8.40	.01	.19	
Externalized Aggression	-.36	.7821	.6116	.0185	4.24	.04	-.63	
Socialization	-.23	.7838	.6143	.0026	.60	.44	.51	
Disrespect/ Defiance	-.37	.7846	.6157	.0014	.32	.58	-.29	
Attention/ Motivation	-.38	.7856	.6171	.0015	.33	.57	.14	
Academic Self-Confidence	-.29	.7867	.6189	.0017	.39	.54	-.34	
Attention-Seeking	-.35	.7872	.6197	.0008	.18	.68	.19	
Sex	.14	.7874	.6200	.0003	.07	.79	.33	

Aggression was subsequently entered adding an additional 6% to the explained variance. Further inclusion of Academic Self-Confidence, and Disrespect-Defiance added an additional 5% to the explained variance in SAT_4. Subsequent inclusion of sex of student after the behavior scales added only negligible amounts to the explained variance. Further inclusion of IQ increased the multiple R to .72 (52% of the total explained variance), and final inclusion of SAT_1 increased the multiple R to .79 (62% of the total explained variance in SAT_4).

Figure 1 graphically summarizes and depicts the relative contribution of teachers' ratings of student behavior to the explained variance in future reading achievement at one to three year intervals, independent from, and in conjunction with initial level of reading achievement, IQ, and sex of student. Accuracy of prediction appeared to be greatly enhanced by knowledge of a student's initial level of achievement, as well as general abilities.

Both IQ and initial level of achievement (SAT_1), independently, proved to be better predictors of achievement than the aggregate of the behavior scales at all time intervals. Attention/Motivation proved to be consistently the best of the behavioral predictors. Figure 2 represents the relative contribution of the behavior scales, IQ, initial level of achievement, and sex of student to the explained variance in future reading achievement at one to three year intervals.

Observation of the relationship between selected behavioral attributes from the identified behavior scales and reading achievement indicated interesting trends in the predictive relationship over time for the total group and boys and girls independently.

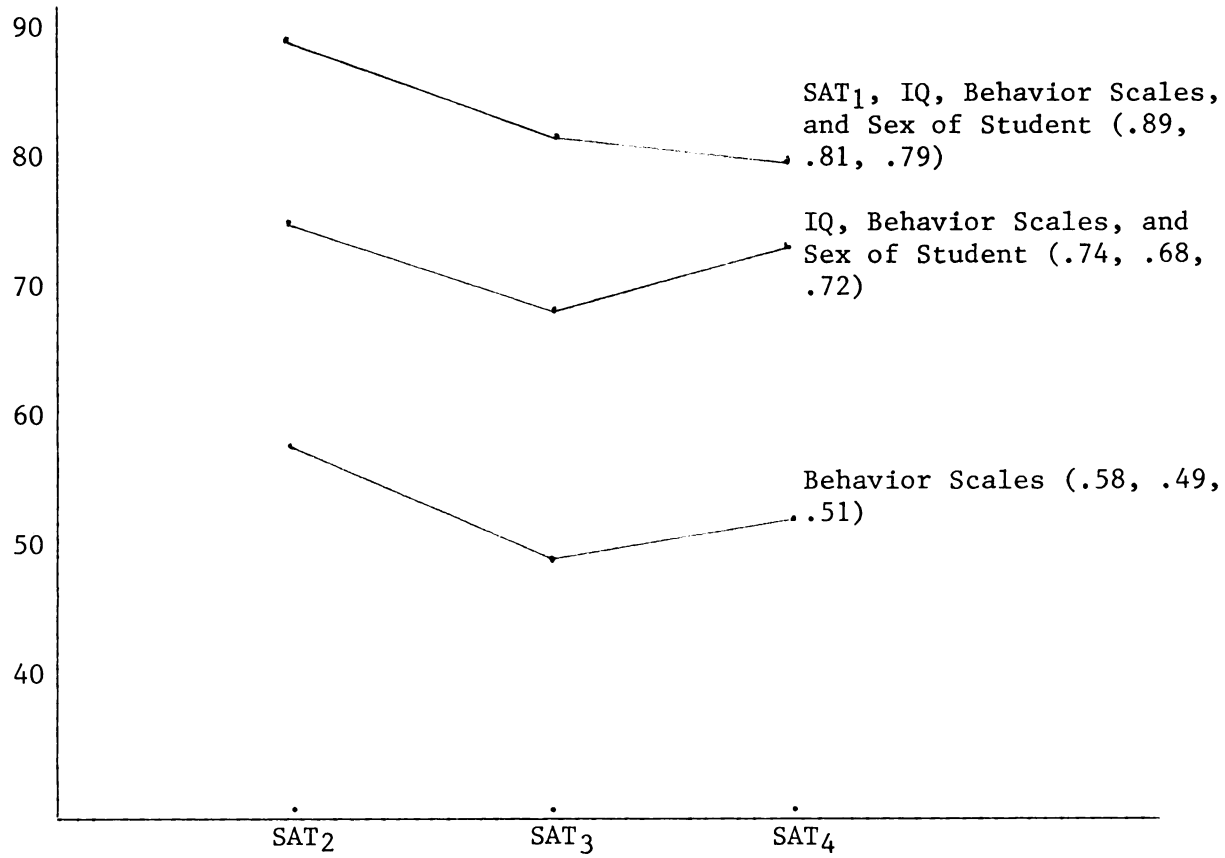

Figure 1. Multiple Correlation of Reading Achievement with a) Behavior Scales, b) Behavior Scales, Sex of Student, and IQ, and c) Behavior Scales, Sex of Student, IQ and SAT_1

Evaluation of selected Attention/Motivation scale items indicated that for the total group, (1) those items which tended to reflect student self-motivation and persistence were highly and significantly related to reading achievement across time (SAT_1 to SAT_4) -- e.g. Tends to give up (-.49, -.44, -.45, -.48), Finds it hard to study (-.42, -.44, -.32, -.35), and Is interested in school work (.38, .38, .29, .30), while, (2) those items which reflected student attending/concentration behaviors tended to decrease in strength and significance of relationship with reading achievement over time -- e.g. Distractible; can't concentrate (-.37, -.40, -.27, -.21), and Able to concentrate on things (-.31, -.35, -.24, -.21).

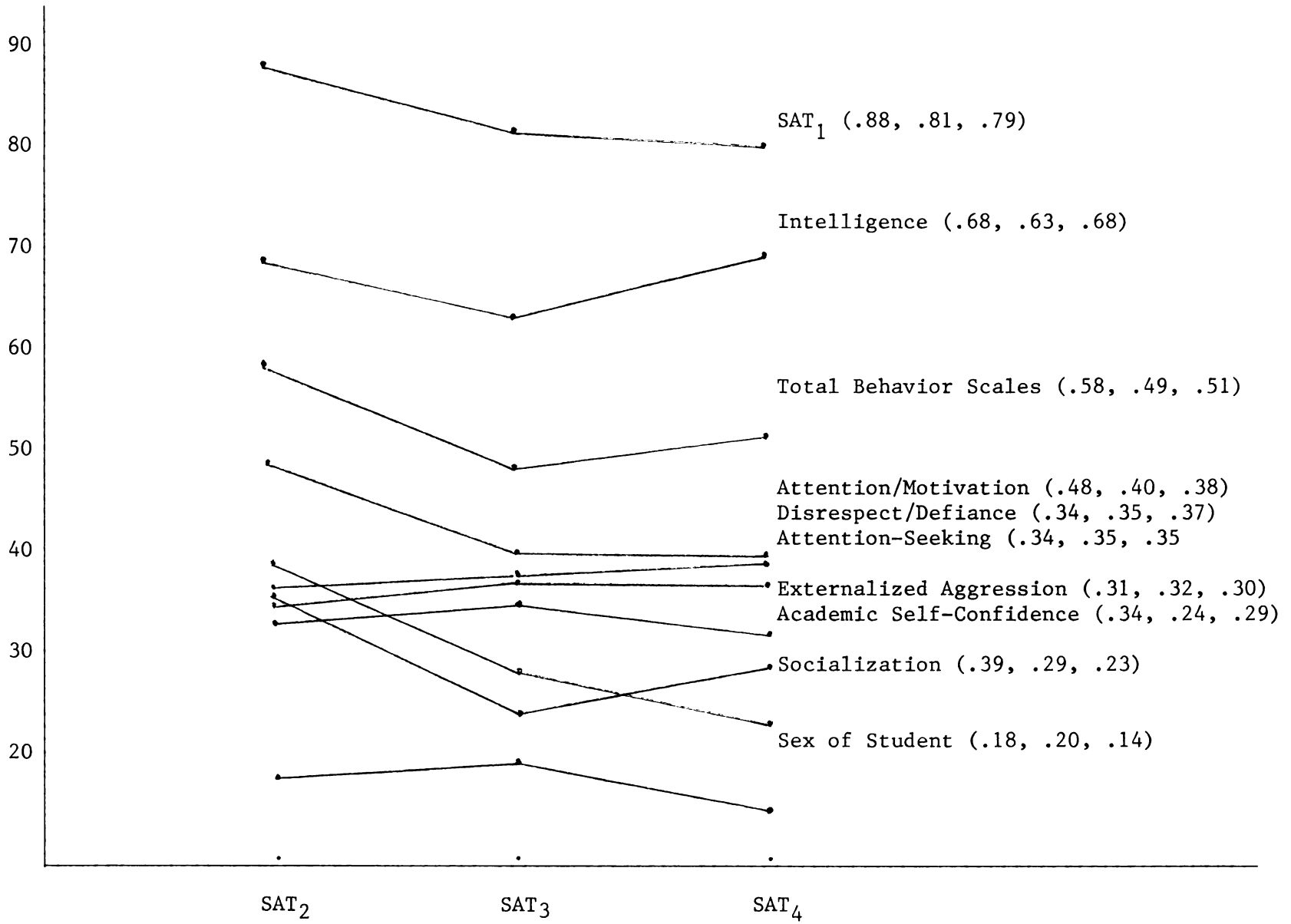

Figure 2. Absolute Correlations of Reading Achievement with the Behavior Scales, IQ, Sex of Student, and SAT_1

Observation of sex differences in the relationship suggested relatively different trends in the relationship of selected Attention/Motivation items without achievement over the four year period. For boys, only three of eight items which were significantly related to SAT_1 were significantly related to SAT_4, and again those items tended to reflect general persistence, motivation, and awareness in the classroom setting -- e.g. Tends to give up (-.49, -.49, -.47, -.46), Finds it hard to study (-.45, -.50, -.28, -.36), and Is alert in class (-.39, -.40, -.31, -.34). For the most part, other items of the Attention/Motivation scale, which were initially highly related to reading achievement for boys for the first two years, tended to decrease sharply in strength and significance over years three and four -- e.g. Finishes classroom assignments (.30, .28, .08, .20), Does homework (.39, .45, .29, .18), Distractible; can't concentrate (-.33, -.36, -.17, -.14), and Lacks ambition to do well in school (-.35, -.33, -.14, -.16). On the contrary for girls, the relationship between initial ratings of Attention/Motivation items and subsequent reading achievement tended to strengthen over time often from initially nonsignificant levels to moderately strong, significant levels -- e.g. "Drags" feet when requested to do something (-.25, -.32, -.30, -.30, -.47), Interested in school work (.21, .32, .27, .36), Lacks ambition to do well (-.18, -.29, -.25, -.35), Does homework (.21, .26, .29, .33), and Finds it hard to study (-.24, -.24, -.30, -.31). Again like their male counterparts, those behaviors found to be highly related to SAT_4 appeared to reflect a generally persistent, motivated orientation to the classroom demands. Similarly, as was the case with the boys, early high correlations between attending/concentrating behaviors tended to decrease to nonsignificant levels by SAT_4 -- e.g. Distractible; can't concentrate (-.37, -.41, -.35, -.26), and Is alert in

class (.32, .37, .39, .14).

Simple correlations between selected Disrespect/Defiance items were generally found to be in the moderately low, nonsignificant range for the total group at SAT_1. Only the item, Sulks when things go wrong, appeared moderately and consistently related to achievement across time for the total group (-.29, -.31, -.28, -.38). The relative strength of correlations appeared to moderately increase over time, with highest correlations between Disrespect/Defiance items evidenced at SAT_4 -- e.g. Will put up an argument when told not to do something (-.25, -.21, -.24, -.26), Gives other children dirty looks (-.16, -.18, -.21, -.32), and Argues with teacher (-.22, -.22, -.24, -.30).

For boys as a group, moderately high correlations across time were consistently evidenced on those Disrespect/Defiance items; Sulks when things go wrong (-.34, -.34, -.36, -.46), Has changeable moods (-.34, -.38, -.41, -.41), and Is infuriated by any form of discipline (-.33, -.29, -.32, -.34). The latter two items (Has changeable moods and Is Infuriated by any form of discipline) evidenced negligible relation to reading achievement for girls at all time samplings, -.04 to .05 and -.13 to .01 respectively. For girls as a group, moderately high correlations were evidenced at SAT_1 and SAT_4 with those Disrespect/Defiance items; When angry will refuse to speak to anyone (-.32, -.18, -.26, -.39), and When angry will do things like slamming the door or banging the desk (-.34, -.26, -.24, -.38).

Correlations appeared relatively consistent across time for the total group on those items of the Attention-Seeking scale. Moderately low relationship between the majority of items and reading achievement were evidenced -- e.g. Deliberately interrupts what is going on by asking

silly questions (-.35, -.32, -.31, -.38), and Tries to be the center of attention (-.32, -.31, -.28, -.33), appeared to be consistently related to reading achievement over the four year period. No items from the Attention-Seeking scale were significant for girls at SAT_1, however, the item, Does things just to attract attention was significantly related to achievement at SAT_4 (-.36).

Moderately low correlations were consistently evidenced between items of the Externalized Aggression scale and reading achievement across time for the total group. The items, Teases other children (-.29, -. 31, -. 27, -.33) and Threatens to hurt other children when angry (-.23, -.28, -.24, -.32) evidenced the strongest relationship with achievement of any items in the scale. Observation of sex differences again, however, evidenced different trends of association between selected items of the Extermalized Aggression scale and reading achievement over time. Despite moderately low correlations with SAT_1, the following items were found to be strongly related to SAT_4 for boys as a group: Threatens to hurt other children when angry (-.47), Teases other children (-.40), Hits and pushes other children (-.40), and Does things to get others angry (-.37). On the contrary, no items from the Externalized Aggression scale were found to be significantly related to reading achievement at any time sampling for girls.

Generally low, yet consistent, patterns of relationship were evidenced across time between items of the Academic Self-Confidence scale and reading achievement for the total group. The item, Seems dull; slow to catch on, was found to be the only item from the scale to be highly and consistently related to achievement across time (-.42, -.44, -.36, -.34). For boys as a group, the aforementioned item again appeared to be consistently highly related to achievement (-.44, -.50, -.37, -.35, while

for girls, that item evidenced only moderate levels of association over time (-.31, -.24, -.27, -.30). The item, Volunteers to recite in class, was significantly related to SAT_2 and SAT_4 for girls (.32 and .31), but found to be not related to reading achievement at any level for boys (.12, .12, -.05, .05).

Hypothesis IV: The linear rule derived from the identified behavior scales does not predict the achievement scores of Good, Average, and Poor readers after one, two and three year intervals.

In order to ascertain the relative efficiency of the identified independent variables (behavior scales) in predicting future reading achievement, the derived equations were evaluated by observing the accuracy of prediction for each student at SAT_2, SAT_3, and SAT_4. The residuals, or differences between each student's attained reading achievement score (Y) and estimated reading achievement score (Y'), were computed. Students were defined as Good, Average, or Poor Readers based on their attained reading achievement score (Y) as well as their predicted reading achievement score (Y') at each year interval and these classifications served as the means by which relative errors in prediction were evaluated. Diagnostic "hits" were evidenced when students' attained and predicted scores fell within one standard error of estimate from each other. Six separate detection tables were then generated, one for each prediction equation at SAT_2, SAT_3, and SAT_4.

Base rates were finally computed for Good, Average, and Poor Readers at each year interval, and conditional probabilities calculated for each of the predictive test signs to determine the relative utility of the predictive equations in forecasting future student reading achievement after one, two, and three years. The conditional probabilities are reflected by the following Baysian expressions:

Probability of Being a Good Reader Given a Good Reader Test Sign

$$P(GR|-) = \frac{P(GR)\ P(-|GR)}{P(GR)\ P(-|GR) + P(PR \text{ or } AR)\ P(-|PR \text{ or } AR)}$$

Probability of Being a Poor Reader Given a Poor Reader Test Sign

$$P(PR|+) = \frac{P(PR)\ P(+|PR)}{P(PR)\ P(+|PR) + P(GR \text{ or } AR)\ P(+|GR \text{ or } AR)}$$

Probability of Being an Average Reader Given an Average Reader Test Sign

$$P(AR|o) = \frac{P(AR)\ P(o\ AR)}{P(AR)\ P(o\ AR) + P(GR \text{ or } PR)\ P(o\ GR \text{ or } PR)}$$

Figure 3 presents the relative predictive accuracy of the behavior scales in terms of correct classifications of Good, Average, and Poor Readers after one year (SAT_2) with respect to the established prediction equation. The behavior scales contributed 33% to the total explained variance in SAT_2(multiple R = .58). Inspection of Figure 3 revealed that the overall "hit rate" in terms of correct classification was 74% (69/93). Eighty percent of the predicted Poor Readers were actually Poor Readers, 74% of the predicted Average Readers were Average, and no Good Readers were predicted by the regression statement. At the extremes, no significant errors in prediction were evidenced, that is, no actual Good Readers were predicted to be Poor Readers, and no actual Poor Readers were predicted to be Good Readers. An extremely high false positive error rate of 100% (0/14) was evidenced in the prediction of Good Readers, while a high false negative rate of 69% (9/13) was evidenced for Poor Readers. Error rate was minimized within the Average group (2%).

The relative predictive accuracy of an equation involving the behavior scales after two years (SAT_3) is presented in Figure 4. With respect to the established prediction equation, the behavior scales contributed 24% to the total explained variance in SAT_3 (multiple R = .49). Figure 4

Predicted \ Actual	Good Reader	Average Reader	Poor Reader	Total
Good Reader	0	0	0	0
Average Reader	14	65	9	88
Poor Reader	0	1	4	5
Total	14	66	13	93

Figure 3. Predictive Classification of Students at SAT_2 Based on the Behavior Scales

indicates that the overall "hit rate" in terms of correct classification was 74.2% (69/93). Seventy-five percent (65/74) of the predicted Average Readers were Average, 100% (1/1) of the predicted Poor Readers were Poor, no Good Readers were predicted by the equation. Again at the extremes, no significant errors in prediction were evidenced, that is, no actual Good Readers were predicted to be Poor Readers, and no actual Poor Readers were predicted to be Good Readers. However, an extremely high false positive error rate of 100% (0/14) was again evidenced in the prediction of Good Readers, while an extremely high false negative rate of 92% (11/12) was evidenced for Poor Readers. Error rate was nil for the average group.

Figure 5 presents the relative predictive accuracy of the behavior scales in terms of correct classification of Good, Average, and Poor Readers after three years (SAT_4). With respect to the established prediction equation the behavior scales contributed 26% to the total explained variance in SAT_4 (multiple R = .51). Inspection of Figure 5 indicated that the overall "hit rate" in terms of correct classification was 75%

Predicted \ Actual	Good Reader	Average Reader	Poor Reader	Total
Good Reader	0	0	0	0
Average Reader	13	68	11	92
Poor Reader	0	0	1	1
Total	13	68	12	93

Figure 4. Predictive Classification of Students at SAT_3 Based on the Behavior Scales

(70/93). Only 50% of those students predicted to be Poor Readers were in in fact Poor Readers, 76% of predicted Average Readers were Average, and no Good Readers were predicted by the regression statement. Again at the extremes, no actual Good Readers were predicted to be Poor Readers. An extreme false negative error rate (100%) was again evidenced in the prediction of Good Readers, while an extreme false positive error rate (.91) was observed in the prediction of Poor Readers utilizing the behavior scales only. Error rate was minimized within the Average group (.01).

Hypothesis V: The linear rule derived from the initial level of reading achievement (SAT_1), IQ, sex of student, and the identified behavior scales does not predict the achievement scores of Good, Average, and Poor readers after one, two and three year intervals.

The relative accuracy of a prediction equation involving the behavior scales, IQ, SAT_1, and sex of student in terms of correct classification of Good, Average, and Poor Readers after one year (SAT_2) is presented in Figure 6. These independent variables contributed 79% to the explained

Predicted \ Actual	Good Reader	Average Reader	Poor Reader	Total
Good Reader	0	0	0	0
Average Reader	12	69	10	91
Poor Reader	0	1	1	2
Total	12	70	11	93

Figure 5. Predictive Classification of Students at SAT_4 Based on Behavior Scales

variance in SAT_3 (multiple R = .89). Evaluation of Figure 6 indicated that 93.5% (87/93) of the students were correctly classified. Relative "hit rates" were high for all groups. One hundred percent of those students predicted to be Good Readers were in fact Good Readers; 92.9% of Average Readers were predicted to be Average; and 91.7% of Poor Readers were predicted to be Poor. Inclusion of IQ, sex, and SAT_1 into the prediction with the behavior scales greatly decreased the false positive error rate to 21% (3/14) for Good Readers, and greatly decreased the false negative error rates for Poor Readers to 15% (2/13).

Figure 7 indicates the relative accuracy of a prediction equation involving the behavior scales, IQ, SAT_1, and sex of student in terms of correct classification of Good, Average, and Poor Readers after two years (SAT_3). Sixty-six percent of the total variance in SAT_3 was explained by the aforementioned independent variables (multiple R = .81). Inspection of Figure 7 revealed that the overall "hit rate" in terms of correct classification was 87% (81/93). Ninety percent of those students predicted

Predicted \ Actual	Good Reader	Average Reader	Poor Reader	Total
Good Reader	11	0	0	11
Average Reader	3	65	2	70
Poor Reader	0	1	11	12
Total	14	66	13	93

Figure 6. Predictive Classification of Students at SAT_2 Based on the Behavior Scales, IQ, SAT_1, and Sex of Student

to be Good Readers were in fact Good Readers; 88% of Average Readers were predicted to be Average; and 78% of Poor Readers were predicted to be Poor Readers. Despite the addition of SAT_1, IQ, and sex of student to the equation in addition to the behavior sclaes, a relatively high false positive rate (.31) persisted in the prediction of Good Readers, as well as a high false negative error rate (.42) in the prediction of Poor Readers.

The relative accuracy of a prediction equation involving the behavior scales, IQ, SAT_1, and sex of student in terms of correct classification of Good, Average and Poor Readers after three years (SAT_4) is presented in Figure 8. Sixty-two percent of the total variance in SAT_4 was explained by the aforementioned independent variables (multiple R = .79). Inspection of Figure 8 indicated an overall "hit rate" of 88% (82/93) in terms of correct classification. Eighty-nine percent of those students predicted to be Good Readers were in fact Good Readers; 88% of Average Readers were predicted to be Average; and 86% of Poor Readers were predicted to be

Predicted \ Actual	Good Reader	Average Reader	Poor Reader	Total
Good Reader	9	1	0	10
Average Reader	4	65	5	74
Poor Reader	0	2	7	9
Total	13	68	12	93

Figure 7. Predictive Classification of Students at SAT_3 Based on the Behavior Scales, IQ, SAT_1, and Sex of Student

Poor Readers. A relatively high false negative error rate (.33) persisted in the prediction of Good Readers, and a similarly high false positive error rate (.45) was evidenced in the prediction of Poor Readers.

Predicted \ Actual	Good Reader	Average Reader	Poor Reader	Total
Good Reader	8	1	0	9
Average Reader	4	68	5	77
Poor Reader	0	1	6	7
Total	12	70	11	93

Figure 8. Predictive Classification of Students at SAT_4 Based on the Behavior Scales, IQ, SAT_1, and Sex of Student

Table 18 summarizes the relative diagnostic "hit rates" by reading group based on 1) the behavior scales only, and 2) the behavior scales, IQ, SAT, and sex of student across time.

TABLE 18

Diagnostic "Hit Rates" of Good, Average, and Poor Readers at SAT_2, SAT_3, and SAT_4

Predictor(s)	SAT_2				SAT_3				SAT_4			
*	T	G	A	P	T	G	A	P	T	G	A	P
Behavior Scales	.74	.00	.73	.80	.74	.00	.75	1.00	.75	.00	.76	.50
Behavior Scales, IQ, SAT_1, Sex	.94	1.00	.93	.92	.87	.90	.88	.78	.88	.89	.88	.86

*T-Total, G-Good, A-Average, P-Poor

As is quite evident, inclusion of SAT_1, IQ, and sex of student in the prediction, greatly enhanced the classification "hit rate" at all levels of achievement. The behavior scales independent of the aforementioned variables provided only marginal predictive accuracy and utility with future reading achievement. A consistent overall hit rate of approximately 75% was observed with most extreme errors in prediction evidenced at the extremes i.e. Good Readers and Poor Readers.

The behavioral scales alone tended to highly underpredict achievement for Good Readers (false positive error rates of 100% at SAT_2, SAT_3 and SAT_4), however, produced no misclassification of students who later became Poor Readers (a more seriuos false positive error). Similarly, the behavior

scales alone tended to highly overpredict achievement for Poor Readers (false negative error rates of 69%, 92%, and 91% for SAT_2, SAT_3, and SAT_4 respectively), however, produced no misclassification of students who later became good readers (an even more serious false negative error).

Overall observed high "hit rates" in prediction with the inclusion of SAT_1, IQ, and sex of student with the behavior scales, again clearly need to be observed relative to the degree of prediction error at the extremes with Good and Poor Readers. False positive errors in predicting Good Readers, and false negative errors in predicting Poor Readers again appeared prevalent, even when using prediction equations with multiple correlations as high as .89. For SAT_2, the equation underpredicted achievement for 21% (3/14) of Good Readers and overpredicted achievement for 15% (2/13) of Poor Readers. For SAT_3, the equation underpredicted achievement for 31% (4/13) of Good Readera, and overpredicted achievement for 42% (5/12) of Poor Readers.

Calculated conditional probabilities of decision risk again suggested relative caution in interpreting the practical utility of the generated prediction equations, especially in terms of decision-making for individual students whose predicted scores fell at the extreme ranges, i.e. Good Reader or Poor Reader.

When only the behavior scales were incorporated in the prediction of future reading achievement the following conditional probabilities were attained for each of the test prediction signs: Poor Reader (+), Average Reader (0), Good Reader (-). For SAT_2, SAT_3, and SAT_4, the probabilities that a student would be a Poor Reader given Poor Reader predicted scores P(PR +) were .29, .07, and .08 respectively. Again, for SAT_2, SAT_3, and SAT_4, the probabilities that a student would be a Good Reader given Good Reader predicted scores P(GR -) were .00 at all three intervals. Finally,

for SAT_2, SAT_3, and SAT_4, the probabilities that a student would be an Average Reader given Average Reader predicted scores P)AR o) were .99, 1.00, and .99 respectively. These results again suggest the relative instability of the prediction equation with the extreme groups.

When SAT_1, IQ, and sex of student were incorporated with the behavior scales in the prediction of future reading achievement the following conditional probabilities were obtained for each of the test prediction signs. For SAT_2, SAT_3, and SAT_4, the probabilities that a student would be a Poor Reader given Poor Reader predicted scores P(PR +) were .84, .60, and .53 respectively. Again for SAT_2, SAT_3, and SAT_4, the probabilities that a student would be a Good Reader given Good Reader predicted scores P(GR -) were .80, .27, and .69. Finally, for SAT_2, SAT_3, and SAT_4, the probabilities that a student would be an Average Reader given Average Reader predicted scores P(AR o) were all .99.

Results of this final aspect of the study again indicated that:

1) the generated prediction equations which utilized only the behavioral scales in predicted SAT_2, SAT_3, and SAT_4 tended to greatly underpredict reading achievement for Good Readers, to greatly overpredict reading achievement for Poor Readers, and to accurately predict reading achievement for Average Readers.
2) despite significantly high multiple correlations, as well as high "hit rates", the generated prediction equations which incorporated the behavior scales, SAT_1, IQ, and sex of student were further found to be reasonably lacking in practical utility in terms of accurately identifying Good and Poor Readers over time.

CHAPTER V

DISCUSSION

Student behavioral attributes which appeared most highly and significantly related to reading achievement in the present study appeared to generally reflect a student's degree of adaptation to the task orientation of the classroom, ability to attend to classroom tasks, degree of self-motivation and persistence, and degree of general social maturity. The general portrait of the academically competent student appeared similar to previously reported findings (Cobb, 1972; Lambert and Nicoll, 1977; Lahaderne, 1968; McKinney et al., 1975; Samuels and Turnure, 1974). Results suggested that the student who is attentive, interested, confident, persistent, and who regularly completes assignments in the classroom setting is more likely to be a competent reader than the student who is distractible, dependent on the teacher for direction, socially immature, poorly self-motivated, and who does not regularly complete assigned tasks.

At a global level, the identified achievement-related behavior characteristics in the present study appears conceptually consistent with the achievement-related behavioral dimensions defined by Lambert and Nicoll (1977), namely Classroom Adaptation (e.g. following directions, volunteering, working independently), Intrapersonal Adjustment (e.g. withdrawn, attention to task), Interpersonal Adjustment (e.g. fighting, arguing). The behavioral attributes, which appeared to generally reflect the intrapersonal adjustment dimension as defined by Lambert and Nicoll, tended to be most highly related to reading achievement in the present study, (e.g. Tends to give up if has something hard to finish, Is alert in class, Is sure of self, Is interested in schoolwork, and Is distractible; can't concentrate).

Behavioral attributes which reflected adaptation to classroom task expectations were similarly highly related to reading achievement (e.g. Does homework, and Finishes assignments). Volunteers to recite, unlike the findings of Cobb (1972) and Lambert and Nicoll (1977), was only marginally associated with reading achievement in the present study.

As was the case in the studies of Lambert and Nicoll (1977) and Lambert and Urbanski (1980), behaviors reflecting an interpersonal malajustment dimension tended to evidence generally weaker relationships to reading achievement, than those reflecting intrapersonal adjustment and adaptation to classroom demands.

Behaviors reflecting a general socially immature, attention-seeking orientation (e.g. Likes an audience all the time, Deliberately interrupts what is going on by asking silly questions, Does things just to attract attention, and Does things which are normal for children much younger) proved to be more highly related to reading achievement, than those behaviors which indicated a more aggressive, disruptive, acting out posture (e.g. Acts up when adults are not watching, Hits and pushes other children, Threatens to hurt other children when angry, Seems unconcerned when misbehaving, and Disturbs other children with boisterous behavior). For the total group those attributes from the Attention/Motivation Scale which reflect student self-motivation and persistence were found to be highly and consistently related to achievement across time (e.g. Tends to give up, Finds it hard to study, and is interested in schoolwork), wherein those items which reflected student attending/concentration behavior tended to decrease in relative strength and significance over time. The latter finding appears most relevant in that a great amount of past research has reported "attending" to be generally the most significantly related behavioral attribute to achievement for elementary age students

(Cobb, 1970, 1972; Feshbach, et al., 1977; Forness, et al., 1977; Meyers, et al., 1968; Soli and Devine, 1976). Results of the present study suggested a general trend indicating either improved attending abilities over time for elementary age students with consistent reading achievement scores, or relatively consistent gain in reading achievement independent of attending abilities. The relationship at this time is unclear as behavioral data was unable to be attained at each successive time sampling. Explanation of the relationship might be enhanced should relative stability of the behavior ratings be observed.

As was the case in the studies of Cobb (1970), Feshbach, et al. (1977), Samuels and Turnure (1974), and Swift and Spivack (1967), general classroom behavioral adjustment appeared to be a generally more important factor in terms of the reading success of elementary age boys relative to their female peers. What appeared to be the most striking aspect of the observed behavior-achievement relationships for boys versus girls was the moderately strong negative relationships for boys but virtually negligible relationship for girls between reading achievement and those behavioral attributes which reflect a student's degree of lack of behavioral self-control, emotional instability and social immaturity -- behaviors most likely to be observed as disruptive in the classroom setting, (e.g. Likes an audience all the time, Deliberately interrupts what is going on by asking silly questions, Is infuriated by any form of discipline, Interrupts whoever is speaking, and Approaches a task with an air of defeatism).

Different trends in the relationship of the Attention/Motivation items and reading achievement over time were as expected observed between girls and boys. Attributes which were highly related to reading achievement for boys for the first two years, tended to decrease sharply in strength and significance over years three and four. On the contrary for girls, the

relationship between initial ratings of Attention/Motivation items tended to strengthen over time often from initially non-significant levels to moderately strong, significant levels. Such findings suggest that there may be a developmental adjustment phenomenon occuring wherein classroom adaptation skills and attributes become refined over time for boys as a group, and cognitive ability and past achievement may be more important contributing factors to future reading success.

The fact that intrapersonal/motivation variables appeared to be the most highly related to measured reading achievement for boys as a group over time, suggested that potential remedial interventions for elementary boys who are experiencing reading difficulties should be closely scrutinized in terms of their novel, motivational, and "attention-getting" characteristics. External reinforcers such as tangible rewards and high levels of verbal feedback and teacher praise may further be necessary components of developmental or remedial reading programs for elementary age boys.

Moderately low correlations were consistently evidenced between items of the Externalized Aggression scale and reading achievement across time for the total group. Sex differences, however, were again observed in the behavior-achievement relationship with items from the scale. Despite moderately low correlation at initial level of achievement, items such as Threatens to hurt other children when angry, Teases other children, and Hits and pushes other children were found to be strongly associated with reading achievement for boys after four years. On the contrary, no items from the Externalized Aggression scale were found to be significantly related to reading achievement at any time sampling for girls. Studies, such as those of Lambert and Nicoll (1977), Lambert and Urbanski (1980), and Cobb (1972), have suggested the relative lack of relationship between

such externalizing, aggressive behaviors and early reading achievement. Results of the present study suggest that at least for boys, such initially-observed behaviors do indeed relate to achievement over time, and as such should be considered for intervention concurrently with developmental or remedial reading instruction.

Results of past longitudinal studies which have incorporated behavioral attributes do make a significant, unique contribution to the variance in later reading ability, independent of various cognitive measures, including intelligence (Feshbach, et al., 1977; Meyers, et al., 1968; Stevenson, et al., 1976b). Current attained multiple correlations between reading achievement and the aggregate of the behavior scales, appeared highly commensurate with the attained multiple R's attained between the Student Rating Scale (SRS) and reading achievement as measured by the Cooperative Primary Test over a three year period -- multiple R= .50 (Feshbach, et al., 1977).

Reflective of the findings of Stevenson, et al., (1976b), the results of the present study found that despite the observed strong predictive relationship between the behavior scales and future reading achievement, teacher's ratings were not observed to be as predictive of later achievement as were scores on a measure of general intelligence, as well as scores on an initial reading achievement measure. These results would suggest that very little predictive information is gained by knowledge of behavioral attributes of elementary-age students beyond knowledge of initial levels of academic functioning and general intellectual ability as measured by group-administered standardized measures. Economics would dictate, especially at a gross screening level, that attaining behavioral data would probably not be worth the additional cost and effort involved. However, for purposes of prediction for individual students such behavioral

information may be useful for more accurately assessing future achievement.

The relative efficiency and utility of the derived prediction equations were finally evaluated in terms of their ability to accurately predict reading status. Results indicated that the generated prediction equations which utilized only the behavior scales to predict subsequent reading achievement tended to greatly underpredict achievement group for Good Readers, to greatly overpredict reading achievement for Poor Readers, and to accurately predict achievement for Average Readers. Further, despite significantly high multiple correlation, as well as high "hit rates", the generated prediction equations which incorporated initial level of achievement, IQ, and sex of student were found to be generally lacking in terms of accurately predicting Good and Poor Readers over time. Utilization of the derived formulae over the long run tended to screen out a significant number of students who would have been in need of possible reading support services and tended to similarly screen out those students with above average reading capabilities who might have benefited from possible enrichment activities.

The relative "cost" in terms of the number and types of errors evidenced is somewhat difficult to assess, and is highly dependent upon how one views the cost to students in terms of access to appropriate services, as well as fiscal cost to the school district. If a preventative or early intervention program were to be established for students with "high-risk" test signs (i.e. predicted poor achievement) and high false positive rates in prediction were evidenced, the sum result would represent unnecessary expense and effort to the school, as well as potentially serious implications for the student who may be incorrectly "labled" and unnecessarily targeted into "lower track," remedial programs. On the other hand, it would seem that the more potentially detrimental error in terms of long term

implications for the individual student, the school, and community as a whole may be the false negative rate, wherein large numbers of students may be predicted to do well, when in fact the contrary is true.

Evaluation of the results of the present study would suggest that the derived prediction equation which incorporated initial level of achievement, IQ, sex of student, and the behavior scales would be technically adequate for use as a screening tool for potential enrichment or remedial intervention. Despite the potential of the equation as a possible screening tool, the number and type of errors evidenced would suggest that it would not be acceptable as a single primary evaluation tool. As such further individual assessment should be initiated prior to any decision-making regarding an individual student.

CHAPTER VI

SUMMARY AND RECOMMENDATIONS

To reiterate, the purpose of the study was essentially two-fold: 1) to identify teachers' perceptions of behavior attributes of elementary age students which were highly and significantly related to reading achievement over time and 2) to investigate the efficiency and utility of using teachers' ratings of student behavioral attributes to predict future reading performance. A descriptive, longitudinal design was instigated whereby teachers' evaluations of student behaviors were evaluated independently as well as in conjunction with IQ, initial level of reading achievement, and sex of student, as to their ability to predict subsequent reading achievement measured after one, two and three year intervals.

Initially, items from the School Behavior Checklist (Miller, 1972) which were found to be significantly related to test scores on the Stanford Reading Achievement Test were identified. Scales defining student behavior attributes were subsequently later developed incorporating these achievement-related items from the SBCL. Those behavior scales were defined as reflecting the dimensions Attention/Motivation, Disrespect/Defiance, Externalized Aggression, Academic Self-Confidence, and Socialization. These identified behavior scales were next incorporated into prediction equations so as to evaluate the accuracy of using teacher ratings of student behavioral characteristics to predict future reading achievement measured after one, two, and three year intervals, independently and in conjunction with previously identified, relevant moderating variables including initial level of reading achievement, general intelligence, and sex of student. Finally, the practical, functional utility of the derived equations in

predicting future reading achievement was evaluated by observing the relative error in prediction for each student across time. Predictive "hits" and "misses" were calculated for Good, Average, and Poor Readers, and based upon the observed base rates of reading ability or disability, conditional probability values were calculated to again help to determine the accuracy of prediction as well as the practical utility of using the derived formulae as a basis for intervention.

Student behavioral attributes which appeared most highly and significantly related to reading achievement in the present study appeared to reflect a student's degree of adaptation to the task orientation of the classroom, ability to attend to classroom tasks, degree of self-motivation and persistence, and degree of general social maturity. General classroom behavioral adjustment was found to be a generally more important factor in terms of the reading success of elementary age boys relative to their female peers. What appeared to be the most striking aspect of the observed behavior-achievement relationships for boys versus girls was the moderately strong negative relationship for boys but virtually negligible relationship for girls between reading achievement and those behavioral attributes which reflect a student's degree of lack of behavioral self-control, emotional instability and social immaturity -- behaviors most likely to be observed as disruptive in the classroom setting.

All of the derived behavior scales were found to be significantly related to reading achievement after one to three year intervals. The Attention/Motivation scale, proved to be the most predictive of reading achievement at the initial as well as subsequent one year intervals. IQ was found to be significantly related to all of the behavior scales with the exception of the Externalized Aggression scale, and was as might be expected highly and significantly related to reading achievement at all

four year intervals. Sex of student was found to be significantly related to the Attention/Motivation and Attention-Seeking scales.

Generalizable trends were evidenced between items from the generated behavior scales and reading achievement over the four year period. For the total group those items from the Attention/Motivation scale which reflect student self-motivation and persistence were found to be highly and consistently related to achievement across time, wherein those items which reflected student attending/concentration behavior tended to decrease in relative strength and significance over time.

Different trends in the relationship of the Attention/Motivation items and reading achievement over time were as expected observed between boys and girls. Items which were highly related to reading achievement for boys for the first two years, tended to decrease sharply in strength and significance over years three and four. On the contrary for girls, the relationship between initial ratings of Attention/Motivation items tended to strengthen over time often from initially non-significant levels to moderately strong, significant levels. Like their male peers, those behaviors found to be highly related to reading achievement after four years appeared to reflect a generally persistent, motivated orientation to classroom demands.

Most Disrespect/Defiance items reflected moderately low, non-significant relationships with initial level of reading achievement for the total group with only moderate increases in strength of relationship over time. For boys, but not for girls, moderately high correlations with reading achievement were consistently evidenced on those Disrespect/Defiance items which reflect emotional lability and difficulty with authority.

Similarly, items from the Attention-Seeking scale moderately low relationships were observed over time for the total group. For boys,

however, items such as Does things just to attract attention, Deliberately interrups what is going on by asking silly questions, and Tries to be the center of attention appeared consistently related to achievement over the four year period. No items from the Attention-Seeking scale were significantly related to initial achievement for girls.

Moderately low correlations were consistently evidenced between items of the Externalized Aggression scale and reading achievement across time for the total group. Sex differences, however, were again observed in the behavior-achievement relationship with items from the scale. Despite moderately low correlation at initial level of achievement, Threatens to hurt other children when angry, Teases other children, Hits and pushes other children, and Does things to get others angry, were found to be strongly associated with reading achievement for boys after four years. On the contrary, no items from the Externalized Aggression scale were found to be significantly related to reading achievement at any time sampling for girls.

Generally low, yet consistent patterns of relationship were evidenced across time between items of the Academic Self-Confidence scale and reading achievement for the total group. The item, Seems dull; Slow to catch on was found to be the only item from the scale to be highly and consistently related to achievement across time. That item appeared consistently highly related to achievement for boys, and was found to be only moderately associated with achievement over time for girls. The items, Volunteers to recite in class, was found to be significantly related to achievement after two and four years for girls, but was found to be unrelated to achievement at any level for boys.

The final task of the study was to investigate the relative contribution of teachers' ratings of student behavior to explained variance

in future reading achievement and to assess the accuracy and practical ability of predictions based on the derived behavior scales, independently and in conjunction with moderating variables with known strong predictive relationships with achievement. Results of the present study indicated that despite the observed strong linear relationship between the behavior scales and future reading achievement, teachers' ratings were not found to be as predictive of later achievement as were scores on a measure of general intelligence, as well as scores on the initial reading achievement measure. With the effects of initial level of achievement controlled for, none of the behavior scales were found to be significantly related to reading achievement at any point across the follow-up three year period. When controlling for IQ, the relationship between the behavior scales and future achievement was further diminished to moderately low significant levels to nonsignificant levels. Attention/Motivation after one year, Externalized Aggression after one and three years were the only scales to maintain a significant relationship with achievement while controlling for the effects of IQ.

When initial level of achievement and IQ were entered in the prediction formulae with the behavior scales, initial level of achievement subsumed nearly the total explained variance of achievement (78% of a total of 79%) after one year, 64% of a total 66% of the variance in achievement after two years, and 56% of a total 62% of the variance in achievement after three years. Again, inclusion of the cognitive and prior achievement variables greatly diminished the relative contribution of the behavior scales to the explained variance in reading achievement across time.

The derived regression formulae were finally evaluated in terms of their relative efficiency and utility in terms of appropriate classification

of students as Good, Average, or Poor Readers over the three year follow-up period. Results indicated that the generated prediction equations which utilized only the behavior scales in predicting subsequent reading achievement tended to greatly underpredict achievement group for Good Readers, to greatly overpredict reading achievement for Poor Readers, and to accurately predict achievement for Average Readers. Further, despite significantly high multiple correlations, as well as high "hit rates", the generated prediction equations which incorporated initial level of achievement, IQ, and sex of student were found to be generally lacking in terms of accurately predicting Good and Poor Readers over time. Utilization of the derived formulae over the long run tended to screen out a significant number of students who would have been in need of possible reading support services and tended to similarly screen out those students with above average reading capabilities who might have benefited from enrichment activities.

Limitations and Delimitations of the Study

Limitations

1. Numerous threats to the internal validity of the study in terms of interpretation of the results were evidenced due to the longitudinal nature of the study (Campbell and Stanley, 1966). The accuracy by which future reading achievement status of individual students is predicted will be affected by variables which are unable to be controlled given the current research design. These include history and maturation -- e.g. it is not known which students may have received additional reading support services or what other intervening cognitive, emotional, or environmental variables might be influencing current reading performance; experimental mortality -- the reasons

for subject attrition from the original sample are not known, and it is not known why achievement data may be missing for individual students in the follow-up sample (e.g. untestable, absent, systematically excluded); and, effects of testing -- e.g. the reliability of initial teacher ratings, identified behavioral scale scores, SAT reading achievement scores, and Otis-Lennon IQ scores, as well as individual student reactivity to the testing situations, examiner effects, "test-wiseness" of the students, and changes in the form and level of the SAT.

2. Unfortunately, due to the relatively small sample size, it was not feasible to examine in more depth the predictive behavior-achievement relationship by age and/or grade level, viewing potentially relevant developmental phenomenon in the relationship. Similarly, due to the relatively small sample size, a more indepth analysis of sex of student as an important intervening variable was not possible. As such, possibly unique predictors for boys and girls, and/or younger or older elementary students were unable to be clarified.

3. Again from a developmental perspective, it would have been meaningful to ascertain changes in teacher perceptions of student behavior over the course of the follow-up period to further clarify the behavior-achievement relationship for individual students over that three year period.

4. Because relatively high levels of multicollinearity were observed among the behavior scales, the unique contribution of the individual behavior scales was not able to be observed. Further, because of the observed high correlations among initial level of reading achievement, IQ, and the behavior scales, future prediction of reading achievement

might have been enhanced had the independent variables proven to be more orthogonal.

5. The assigned cut-off criteria for Good and Poor Readers (plus or minus one standard deviation) appears relatively stringent as compared to past predictive studies. Accuracy of correct prediction or classification of group membership is strongly related to the arbitrary nature of the cut-off scores selected. Predictive accuracy may have been enhanced with more liberally established cut-off criteria.

Delimitations

1. Due to the nonrandom selection of subjects, generalization of the results of this study are limited to populations of students who compare similarly on the dimensions of age, sex, IQ, SES, race, school characteristics, and geographical region.

2. A major limitation of multiple regression procedures is the relative instability of the assigned regression weights of the predictor variables when applied to subjects in a new sample. Therefore, prediction of reading achievement of students not included in the sample upon which the prediction equations were developed, would not be meaningful. Cross-validation procedures, that is, means by which the regression weights could be validated, were unable to be implemented because of the relatively small number of subjects in the present study.

Assumptions

In light of the aforementioned limitations and delimitations, the following assumptions are made:

1. The measurement tools employed are reliable indicators of student status (reading achievement, intelligence, and behavioral attributes).

2. Students' performance on the reading and intelligence measures approximate their actual classroom performance and abilities.

3. Measurement errors as a result of history, maturation, experimental mortality, and testing are randomly distributed.

4. For purposes of prediction, the criterion values (reading achievement scores) are randomly distributed.

5. Demographic characteristics of the follow-up sample are representative of the school and district at large from which it was taken.

Recommendations

In light of the observed results and limitations of the present study of prediction of reading achievement over time, the following suggestions are made for future research:

1. Predictive accuracy from a longitudinal perspective might be significantly enhanced by using different behavioral predictors for boys and girls than just including sex as an independent variable in the prediction equation.

2. In order to further clarify the observed trends in the behavior/achievement relationship over time in the present study, future longitudinal studies should attempt to obtain concurrent ratings of student behavior with reading achievement at each time sampling.

3. As initial level of reading achievement, was highly and significantly related to future achievement levels, cummulative achievement scores might enhance the predictive utility of the equations, that is raising multiple correlation above .90 and minimizing errors in prediction for individual students.

4. Cross validation and/or replication studies should be instigated to evaluate the prediction equations' ability to predict achievement of students other than those in the sample on which the formulae were developed.

REFERENCES

Achenbach, T.M., and Edelbrock, C.S. Behavioral problems and competencies reported by parents of normal and disturbed children aged four through sixteen. Monographs of the Society for Research in Child Development, 1981, 46(1), Serial No. 188.

Adelman, H. and Feshbach, S. Predicting reading failure: beyond the readiness model. Exceptional Children, 1971, 37, 349-54.

Anderson, I., Hughes, B., and Dixon, W. The rate of reading development and its relation to learning to read, sex, intelligence, and reading achievement in the sixth grade. Journal of Educational Research, 1956, 49, 447-53.

Anderson, I., Hughes, B., and Dixon, W. The rate of reading development and its relation to age of learning to read, sex, and intelligence. Journal of Educational Research, 1957, 50, 481-94.

Anthony, J. Behavior disorders. In P. Mussen (Ed.), Charmichael's Manual of Child Psychology. New York: Wiley, 1970.

Ashcraft, C. Later school achievement of treated and untreated emotionally handicapped children. Proceedings of the 78th Annual Convention of the American Psychological Association, 1970, 5, 651-52.

Ayllon, T. and Roberts, M. Eliminating discipline problems by strengthening academic performance. Journal of Applied Behavioral Analysis, 1974, 7, 71-76.

Ayres, L. Laggards in Our Schools. New York: Russell Sage Foundation, 1909.

Bannatyne, A. Language, Reading, and Learning Disabilities. Springfield, Ill: Charles C. Thomas, Publisher, 1971.

Bateman, B. Learning disorders. Review of Educational Research. 1966, 26, 93-119.

Bentzen, F. Sex ratios in learning and behavior disorders. National Elementary Principal, 1966, 46(2), 13-17.

Berk, L., Rose, M. and Stewart, D. Attitudes of English and American children toward their school experience. Journal of Educational Psychology, 1970, 61, 33-40.

Blom, G.E. Sex differences in reading disabilities. In E. Calkins (Ed.), Reading Forum, NINDS Monograph No. 11. Washington, D.C.: U.S. Printing Office, 1970, 31-46.

Blom, G.E., Frey, W.D., Prawat, R.S., and Jarvis, R.L. Sex Differences in Receptive Language and Behavioral Expression as They Relate to Early Reading Achievement. Unpublished manuscript, East Lansing, MI: Michigan State University, 1980.

Bower, E.M. and Holmes, J.A. Emotional factors and academic achievement. Review of Educational Research, 1959, 29, 529-44.

Camp, B.W. and Zimet, S.G. The relationship of teacher rating scales to behavior observations and reading achievement of first grade children. Journal of Special Education, Winter 1974, 8, 353-59.

Carroll, M.W. Sex differences in reading readiness at the first grade level. Elementary English, 1948, 25, 370-75.

Chall, J.S. and Feldmann, S.C. A Study in Depth of First Grade Reading. Cooperative Research Project No. 2728, USOE, 1966.

Chamberlin, R.W. The use of teacher checklists to identify children at risk for later behavioral and emotional problems. American Journal of Diseases of Children, 1976, 130, 141-45.

Clark, M. Reading Difficulties in Schools. Harmondsworth: Penguin, 1970.

Cobb, J.A. The Relationship of Observable Classroom Behaviors to Achievement of Fourth Grade Pupils. Unpublished doctoral dissertation, University of Oregon, 1969.

Cobb, J.A. Survival Skills and First Grade Academic Achievement (Technical Report No. 1). Eugene, OR: Center at Oregon for Research in the Behavioral Education of the Handicapped, December 1970.

Cobb, J.A. Relationship of discrete classroom behaviors to fourth grade academic achievement. Journal of Educational Psychology, 1972, 63, 74-78. (a)

Cobb, J.A. and Hops, H. Effects of academic survival skills in training on low achieving first graders. The Journal of Educational Research, 1973, 67, 108-13.

Cobb, J.A. Academic survival skills and elementary academic achievement. In E. Meyen, G. Vergason, and R. Whelan (Eds.) Strategies for Teaching Exceptional Children. Denver: Love, 1972. (b)

Coker, H. and Lorentz, J. Growth in reading as a correlate of pupil classroom behavior. In P. David Pearson (Ed.) Reading: Theory, Research and Practice. Twenty-sixth Yearbook of the National Reading Conference, 1977, 118-25.

Commins, W.D. More about sex differences. Schools and Society, 1928, 28, 599-600.

Duffy, O., Clair, T., Egeland, B., and Dinello, M. Relationship of intelligence, visual-motor skills, and pyscholinguistic abilities with achievement in the third, fourth, and fifth grades: a follow-up study. Journal of Educational Psychology, 1972, 63, 358-62.

Dykstra, R. and Tinney, R. Sex differences in reading readiness, first grade achievement and second grade achievement. In J.A. Figurel (Ed.), Reading and Realism. Newark Del: International Reading Association, 1969.

Eme, R.F. Sex differences in childhood psychopathology: a review. Psychological Bulletin, 1979, 86, 574-95.

Engelmann, S. Relationship between psychological theories and the act of teaching. Journal of School Psychology, 1967, 2, 93-100.

Engin, A. Prediction of classroom achievement using intelligence and behavioral variables for inner-city children. Psychological Reports, 1975, 36, 67-76.

Entwisle, D.R. Implication of language socialization for reading models and for learning to read. Reading Research Quarterly, 1971, 7, 11-67.

Feshbach, S., Adelman, S., and Fuller, W. Early identification of children with high risk of reading failure. Journal of Learning Disabilities, 1974, 7, 49-54.

Feshbach, S., Adelman, H., and Fuller, W. Prediction of reading and related academic problems. Journal of Educational Psychology, 1977, 69(4), 299-308.

Forness, S., Hall, R., and Guthrie, D. Eventual school placement of kindergartners observed as high risk in the classroom. Psychology in the Schools, 1977, 14, 315-17.

Friedman, R., Fuerth, J.H., and Forsythe, A. A brief screening battery for predicting school achievement at ages seven and nine years. Psychology in the Schools, 1980, 17, 340-46.

Fry, E.B. First Grade Reading Instruction Using a Diacritical Marking System, the ITA, and a Basal Reading System. Cooperative Research Project No. 2745, USOE, 1965.

Gates, A.I. The role of personality maladjustment in reading disability. Pedagogical Seminary, 1941, 59, 77-83.

Gates, A.I. Sex differences in reading abililty. Elementary School Journal, 1961, 61, 431-34.

Gates, A.I. The role of personality maladjustment in reading disability. In G. Natchez (Ed.), Children with Reading Problems. New York: Basic Books, 1968.

Gates, A.I. and MacGinitie, W.H. Technical Manual for the Gates-MacGinitie Reading Tests. New York: Teachers College Press, Columbia University, 1965.

Gentile, L.M. and McMillan, M.M. When Johnny can't read but Mary can, men can help. The Reading Teacher, 1976, 29, 771-75.

Gibson, E.J. and Levin, H. The Psychology of Reading. Cambridge, MA: MIT Press, 1975.

Gilbert, G.M. A survey of referral problems in metropolitan child guidance centers. A Journal of Clinical Psychology, 1957, 13, 37-42.

Glavin, J.P. Persistence of behavior disorders in children. Exceptional Children, 1972, 38, 367-76.

Gordon, R.A. Issues in multiple regression. American Journal of Sociology, 1968, 73, 592-616.

Gove, W. and Herb, T. Stress and mental illness among the young: a comparison of the sexes. Social Forces, 1974, 53, 256-65.

Goyette, C.H., Conners, C.K., and Ulrich, R.F. Normative data on revised Conners Parent and Teacher Rating Scales. Journal of Abnormal Child Psychology, 1978, 6(2), 221-36.

Hahn, H.T. A Study of the Relative Effectiveness of Three Methods of Teacher Reading in Grade One. Cooperative Research Project No. 2687, USOE, 1965.

Hall, V., Hppertz, J., and Levi, A. Attention and achievement exhibited by middle- and lower-class black and white elementary school boys. Journal of Educational Psychology, 1977, 69(2), 115-20.

Hall, R.C., Lund, D. and Jackson, D. Effects of teacher attention on study behavior. Journal of Applied Behavioral Analysis, 1968, 1, 1-12.

Harrison, P. Intelligence and Classroom Behavior as Predictors of Achievement in the Underachieving Child. Unpublished doctoral dissertation, University of Tennessee, 1975.

Heilman, A.W. Sex differences in intellectual abilities. Journal of Educational Psychology, 1933, 24, 47-62.

Heilman, A.W. Effects of Intensive In-Service Program on Teachers' Classroom Behavior and Pupils' Reading Achievement. Cooperative Research Project No. 2709, USOE, 1965.

Heilman, A.W. Prinicples and Practices of Teaching Reading. Columbus, Ohio: Charles E. Merrill Books, Inc., 1967.

Hewett, F.M., Taylor, F.D., and Artuso, A.A. The Santa Monica Project: Demonstration and Evaluation of an Engineered Classroom Design with Emotionally Disturbed Children in the Public Schools, Phase I: Elementary Level. Final Report, Project No. 62893, USOE, DHEW, 1967.

Hoge, R.D. and Luce, S. Predicting academic achievement from classroom behavior. Review of Educational Research, Summer 1979, 49(3), 479-96.

Holmes, J.A. Personality characteristics of the disabled reader. Journal of Developmental Reading, 1961, 111-22.

Hops, H. The school psychologist as a behavior management consultant in a special class setting. Journal of School Psychology, 1971, 9(4).

Hops, H. and Cobb, J.A. Initial investigations into academic survival-skill training, direct instruction, and first grade achievement. Journal of Educational Psychology, 1974, 66, 548-53.

Horton, T.R. The Reading Standards of Children in Wales. Slough: NFER, 1973.

Hughes, M.C. Sex differences in reading achievement in the elementary grades. In Clinical Studies in Reading II: Supplementary Educational Monographs No. 77, 1953.

Johnson, D.D. Sex differences in reading across cultures. Reading Research Quarterly, 1973-74, 9(1), 67-86.

Katz, M. Personal communication. State of Michigan: Department of Education, February 27, 1980.

Kellam, S.G., Branch, J.D., Agrawal, K.C., and Ensminger, M.E. Mental Health and Going to School. Chicago: University of Chicago Press, 1975.

Kellmer Pringle, M.L., Butler, N.R., and Davie, R. 11,000 Seven-Year-Olds. First Report of the National Child Development Study. London: Longmans, 1966.

Kerlinger, F.N. and Pedhazur, E.J. Multiple Regression in Behavioral Research. New York: Holt, Rinehart, and Winston, Inc., 1973.

Kessler, J. Psychopathology of Childhood. Englewood Cliffs, N.J.: Prentice-Hall, 1966.

Khan, S.B. Affective correlates of academic achievement. Journal of Educational Psychology, 1969, 60, 216-21.

Kim, Jr. and Kohout, F. Multiple regression analysis: Subprogram regression. In Nie, H., Hull, C., Jenkins, J., Steinbrenner, K., and Bent, D. (Eds.), Statistical Package for the Social Sciences, Second Edition. New York: McGraw-Hill Book Company, 1975.

Kim, Y., Anderson, H.E., and Bashaw, W.L. The simple structure of social maturity at the second grade level. Educational and Psychological Measurement, 1968, 28, 145-53.

Klein, A. The validity of the screening test of academic readiness in predicting achievement in first and second grades. Educational and Psychological Measurement, 1977, 37, 493-99.

Kohn, M. and Rosman, B. A two factor model of emotional disturbance in the young child: validity and screening efficiency. Journal of Child Psychology and Psychiatry, 1973, 14, 31-56.

Kohn, M. and Rosman, B. Social emotional, cognitive, and demographic determinants of poor school achievement: implications for a strategy of intervention. Journal of Educational Psychology, 1974, 66, 267-76.

Lahaderne, H.M. Attitudinal and intellectual correlates of attention: a study of four sixth grade classrooms. Journal of Educational Psychology, 1968, 59, 320-24.

Lambert, N.M. Intellectual and non-intellectual predictors of high school status. Journal of Special Education, 1972, 6, 247-59.

Lambert, N.M., Hartsough, C.S., and Zimmerman, I.L. Comparative predictive efficiency of intellectual and non-intellectual components of high school functioning. American Journal of Orthopsychiatry, 1976, 46, 109-22.

Lambert, N.M. and Nicoll, R.C. Conceptual model for non-intellectual behavior and its relationship to early reading achievement. Journal of Educational Psychology, 1977, 69(5), 481-90.

Lambert, N. and Urbanski, C. Behavioral profiles of children with different levels of achievement. Journal of School Psychology, 1980, 18, 58-66.

Lampouse, R. and Monk, M. Behavior deviations in a representative sample of children. American Journal of Orthopsychiatry, 1964, 34, 436-46.

Lehrke, R. Sex linkage: a biological basis for greater male vulnerability in intelligence. In R. Osbourne (Ed.), Human Variation. New York: Academic Press, 1978.

Lesser, K. and Bridges, J. The prediction of learning problems in a rural setting: can we improve on readiness tests? Journal of Learning Disabilities, 1973, 6, 90-94.

Lipe, D. and Jung, S.M. Manipulating incentives to enhance school learning. Review of Educational Research, 1971, 41, 249-80.

Luce, S. and Hoge, R. Relations among teacher rankings, pupil-teacher interactions, and academic achievement: a test of the teacher expectancy hypothesis. American Educational Research Journal, Fall 1978, 15(4), 489-500.

Maccoby, E.E. Sex differences in intellectual functioning. In E. Maccoby (Ed.), The Development of Sex Differences, Stanford: Stanford University Press, 1966, 25-55.

Maccoby, E.E. and Jacklin, C.N. The Psychology of Sex Differences. Stanford: Stanford University Press, 1974.

Madden, R., Gardner, R.F., Rudman, H.C., Karlson, B. and Merwin, J.C. The Stanford Achievement Test: Technical Data Report. New York: Harcourt, Brace, Jovanovich, Inc., 1975.

Madsen, C.H., Becker, W.C., and Thomas, D.R. Rules, praise, and ignoring: elements of elementary classroom control. Journal of Applied Bevahorial Analysis, 1968, 1, 139-50.

Manning, J.C. Evaluation of Level-Designed Visual-Auditory and Related Writing Methods of Reading Instruction in Grade One. Cooperative Research Project No. 2650, USOE, 1966.

McCarthy, D. Personality and learning. American Council on Educational Studies, 1949, 13, 93-96.

McFarlane, J., Allen, L. and Honzik, M. A Developmental Study of the Behavior Problems of Normal Children Between the Ages of Twenty-One Months and Fourteen Years. Berkeley: University of California Press, 1954.

McGuide, C., Hindsman, E., King, F.J. and Jennings, E. Dimensions of talented behavior. Educational and Psychological Measurement, 1961, 31, 3-38.

McKinney, J., Mason, J., Perkerson, K. and Clifford, M. Relationship between classroom behavior and academic achievement. Journal of Educational Psychology, 1975, 67, 198-203.

McLaren, V.M. Socioeconomic status and reading ability: a study of infant reading. In Scottish Council for Research in Education, Studies in Reading, Vol. 2. London: University of London Press, 1950.

Meyers, C.E., Attwell, A.A., and Orpet, R.E. Prediction of fifth grade achievement from kindergarten test and rating data. Journal of Educational Measurement, 1968, 23, 457-63.

Miller, L.C. School Behavior Checklist: an inventory of deviant behavior for elementary school children. Journal of Consulting and Clinical Psychology, 1972, 38, 134-44.

Miller, L.C. School Behavior Checklist Manual: Form A1, Ages 3-6 and Form A2, Ages 7-13. Louisville, KY: University of Louisville, 1975.

Morris, J.M. Standards and Progress in Reading. Slough: NFER, 1966.

Munger, R.F., Winkler, R.C., Tergland, J.J., and Kranzler, G.D. Counseling and Guidance for Underachieving Fourth Grade Students. Cooperative Research Project No. 2196, University of North Dakota, 1964.

Murphy, H.A. Reading Achievements in Relation to Growth in Perception of Word Elements in Three Types of Beginning Reading Instruction. Cooperative Research Project No. 2675, USOE, 1965.

Naiden, N. Ratio of boys to girls among disabled readers. The Reading Teacher, 1976, 29(5), 439-42.

Neale, D.C., Gill, N., and Tismer, W. Relationship between attitudes toward school subjects and school achievement. Journal of Educational Research, 1970, 63, 232-37.

Nie, N.H., Hull, C.H., Jenkins, J.G., Steinbrenner, K., and Bent, D.H. Statistical Package for the Social Sciences, 2nd Edition, New York: McGraw-Hill Book Company, 1975.

Nila, Sister M. Foundations of a successful reading program. Education, 1953, 73, 543-55.

Otis, A.S. and Lennon, R.T. Manual for the Otis-Lennon Mental Ability Test. New York: Harcourt, Brace, and Jovanovich, 1967.

Parsley, K.M., Powell, M., O'Connor, H.A. and Deutsch, M. Are there really sex differences in achievement? Journal of Educational Research, 1963, 57, 210-12.

Patterson, G., Littman, R., and Bricker, W. Assertive behavior in children: a step toward a theory of aggression. Monographs of the Society for Research in Child Development, 1967, 32(4), Serial No. 113.

Patterson, G.R., Shaw, D.A. and Ebner, M.J. Teachers, peers, and parents as agents of change in the classroom. In F. Benson (Ed.) Modifying Deviant Social Behaviors in Various Classroom Settings. Eugene, OR: University of Oregon, College of Education, Monograph No. 1, 1969.

Peterson, D.R. Behavior problems of middle childhood. Journal of Consulting Psychology, 1961, 25, 205-09.

Prawat, R. and Jarvis, R. Gender difference as a factor in teachers' perceptions of students. Journal of Educational Psychology, 1980, 72, 743-49.

Preston, R.C. Reading achievement of German and American children. School and Society, 1962, 90, 350-54.

Quay, H. The Structure of Children's Behavior Disorders. Paper read at colloquia at the University of Minnesota, 1965.

Richman, N., Stevenson, J., and Graham, J. Prevalence of behavior problems in three year old children. Journal of Child Psychology and Psychiatry, 1975, 16, 277-87.

Roberts, J. and Baird, J.T. Behavior Patterns of Children in School: United States (Vital and Health Statistics, Series 11). Rockville MDL U.S. Department of HEW, Division of Health Examination Statistics, February 1972.

Rogers, D.G. The C-R Word Recognition Test. Christchurch, N.Z.: Psychological Service, Department of Education, 1971.

Ross, A.O., Lacey, H., and Parton, D. The development of a behavior checklist for boys. Child Development, 1965, 36, 1,013-027.

Rubin, R.A. and Balow, B. Prevalence of teacher identified behavior problems: a longitudinal study. Exceptional Children. 1978, 45(2).

Russell, D.H. Interrelationships of the language arts and personality. Child Development and the Language Arts: Bulletin of the National Conference on Research in English. Champaign, IL: National Council of Teachers of English, 1953.

Rutter, M., Tizard, J., and Whitmore, K. Education, Health and Behavior. London: Longman, 1970.

Rutter, M. Classification. In M. Rutter and L. Hersor (Eds.) Child Psychiatry. Oxford, England: Blackwell Scientific, 1977.

Rutter, M. and Yule, W. Reading difficulties. In M. Rutter and L. Hersor (Eds.), Child Psychiatry. Oxford, England: Blackwell Scientific, 1977.

Samuels, S.J. and Turnure, J.E. Attention and reading achievement in first grade boys and girls. Journal of Educational Psychology, 1974, 66, 29-32.

Satz, P. and Fletcher, J. Early screening tests: some uses and abuses. Journal of Learning Disabilities, 1979, 12, 65-69.

Satz, P. and Friel, J. Some predictive antecedents of specific reading disability: a preliminary two-year follow-up. Journal of Learning Disabilities, 1974, 7, 48-55.

Satz, P. and Friel, J. Predictive validity of an abbreviated screening battery. Journal of Learning Disabilities, 1978, 11, 20-24.

Schlosser, L. and Algozzine, B. The disturbing child: he or she? Alberta Journal of Educational Research, 1979, 25, 30-36.

Schneyder, J.W., Schultz, C.B., and Cowen, S. Comparison of Reading Achievement for First Grade Children Taught by a Linguistic Approach and a Basal Reader Approach. Cooperative Research Project No. 2666, USOE, 1966.

Sheldon, W.D. and Lashinger, D.R. Effects of First Grade Instruction Using Basal Readers, Modified Linguistic Materials, and Linguistic Readers. Cooperative Research Project No. 2683, USOE, 1966.

Slossen, R.L. Slossen Oral Reading Test (SORT). East Aurora, NY: Slossen Educational Publications, 1963.

Soli, S. and Devine, V. Behavioral correlates of achievement: a look at high and low achievers. Journal of Educational Psychology, 1976, 68, 335-41.

Southgate, V. Southgate Group Reading Tests. London: University of London Press, 1959 and 1962.

Spache, G.D. Personality patterns of retarded readers. Journal of Educational Research, 1957, 50, 488-93.

Spache, G.D. Investigating the Issues of Reading Disability. New York: Allyn and Bacon, 1976.

Spache, G.D., Andress, M.C., Curtis, H.A., Rowland, M.L., and Fields, M.H. A Study of a Longitudinal First Grade Reading Readiness Program. Cooperative Research Project No. 2742, USOE, 1965.

Spencer, D.V. and Moquin, L.D. Individualized Reading Versus a Basal Reader Program at the First Grade Level in Rural Communities. Cooperative Research Project No. 2673.

Spivak, G. and Swift, M.S. The Devereux Elementary School Rating Scales: a study of the nature and organization of achievement-related disturbed classroom behavior. Journal of Special Education, 1966, 1, 71-90.

Stanchfield, J.M. Boys' achievement in beginning reading. In J. Figurel (Ed.), Reading and Inquiry. Proceedings of the Annual International Reading Association Conference, 1965, 10, 290-93.

Stanchfield, J.M. Sex Differences in Learning to Read, Bloomington: Phi Delta Kappa Educational Foundation, 1973.

Stevenson, H.W., Parker, T., Wilkinson, A., Hegion, A., and Fish, E. Longitudinal study of individual differences in cognitive development and scholastic achievement. Journal of Educational Psychology, 1976, 68, 507-17.

Stevenson, H.W., Parker, T., Wilkinson, A., Hegion, Al, and Fish, E. Predictive value of teacher ratings of young children. Journal of Educational Psychology, 1976, 68, 507-17.

Stroud, J.B. and Lindquist, E.F. Sex differences in achievement in the elementary and secondary schools. Journal of Educational Psychology, 1942, 33, 657-67.

Swift, M. and Spivack, G. The assessment of achievement related classroom behavior. Journal of Special Education, 1967, 2(2), 137-49.

Thompson, G.B. Sex differences in reading attainments. Educational Research, 1975, 18, 16-23.

Vernon, P.E. The Standardization of a Graded Word List. London: University of London Press, 1938.

Walker, H.M. Walker Problem Behavior Identification Checklist Manual. Western Psychological Services, 1970.

Walker, H.M., Mattson, R.H., and Buckley, N.K. The functional analysis of behavior within an experimental class setting. In W.C. Becker (Ed.) An Empirical Basis for Change in Eduation. Chicago: Science Research Associates, 1971, 236-63.

Werry, J.S. and Quay, H.C. Observing the classroom behavior of elementary school children. Exceptional Children, 1969, 35, 461-70.

Wilf, V. Observable Student Behaviors as Indicators of Emotional Maladjustment and Reading Disability. Unpublished doctoral dissertation, Arizona State University, 1972.

Wozencraft, M. Sex comparisons of certain abilities. Journal of Educational Research, 1963, 57, 21-27.

Young, R. An Evaluation of Elementary Categorical Programs: 1979-80. Lansing, MI: Office of Evaluation Services, Lansing School District, November 30, 1980.

Young-Maston, I. Behavior problems of elementary school children: a descriptive and comparative study. Genetic Psychology, Monographs, 47, 1938, 123-81.

Zax, M., Cohen, E.L., Izzo, L.D. and Trost, M.A. Identifying emotional disturbance in the school settings. American Journal of Orthopsychiatry, 34, 1964, 447-54.

APPENDICES

APPENDIX A

Breakdown of Mean Stanford Reading Achievement Test Standard Scores By Grade By Sex of Student

	SAT_1		SAT_2		SAT_3		SAT_4	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Grade 2	50.6	10.3	50.7	11.0	47.8	9.6	49.8	9.5
Male	49.4	11.3	49.6	12.3	46.7	10.4	49.8	9.7
Female	52.4	9.1	52.3	9.3	50.9	7.7	49.8	10.0
Grade 3	50.4	9.0	52.7	10.2	53.5	9.5	54.3	9.4
Male	44.0	6.9	46.1	8.5	46.9	5.6	46.9	5.9
Female	55.0	7.4	56.9	9.2	57.7	9.2	59.7	7.5
Grade 4	51.2	6.6	52.5	6.3	54.4	6.7	51.5	7.5
Male	51.3	8.0	53.8	7.5	55.7	6.5	52.4	9.2
Female	51.1	5.1	51.3	4.8	53.2	6.8	50.4	5.0
Grade 5	45.8	8.2	48.0	8.3	47.6	7.9	48.1	7.2
Male	44.5	8.8	46.5	9.1	46.9	8.6	48.4	7.5
Female	49.0	6.1	52.2	3.2	49.1	6.5	47.7	7.0
Grade 6	50.9	9.6	49.3	8.4	51.5	8.9	50.5	9.7
Male	52.4	11.7	51.0	10.9	53.0	11.0	51.9	11.1
Female	49.4	7.1	47.5	4.9	50.1	6.2	49.4	8.9
Total	49.7	8,8	50.6	8.9	50.9	8.8	50.8	8.7

APPENDIX B

Breakdown of Mean Behavior Scale Raw Scores
By Grade By Sex of Student

	Attention Motivation		Disrespect Defiance		Attention Seeking		Academic Self-Confidence		Externalized Agression		Socialization	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D	Mean	S.D.
Grade 2	17.6	5.4	12.3	3.1	10.5	2.1	6.6	0.9	9.8	2.5	4.1	1.1
Male	17.8	5.3	12.3	3.1	10.5	2.2	6.8	1.1	9.9	2.6	4.3	1.1
Female	17.3	5.9	12.1	3.2	10.4	2.1	6.1	0.4	9.6	2.5	3.8	1.0
Grade 3	16.9	4.4	12.7	3.1	12.7	3.1	6.3	2.8	9.8	2.8	3.3	0.8
Male	19.9	4.1	14.6	3.4	11.0	3.4	6.0	0.8	11.4	3.0	3.4	0.7
Female	14.7	3.4	11.4	2.0	10.0	1.5	6.5	1.0	8.7	2.1	3.3	0.9
Grade 4	14.4	2.7	11.7	2.3	9.0	2.0	6.8	1.4	8.8	1.9	3.3	0.2
Male	14.8	3.6	11.4	1.9	10.5	2.7	6.9	1.7	9.3	2.3	3.3	0.6
Female	14.0	1.4	12.0	2.7	9.3	0.5	6.6	1.0	8.3	1.2	3.3	0.8
Grade 5	17.4	4.2	13.8	3.1	10.9	2.6	6.8	1.0	9.0	1.6	3.3	0.8
Male	18.3	4.3	13.9	2.9	11.3	2.9	6.6	0.9	9.2	1.8	3.4	0.9
Female	14.3	1.5	13.7	3.7	9.9	1.5	7.0	1.1	8.4	0.8	3.3	0.5
Grade 6	16.5	4.3	12.6	2.8	10.4	2.2	7.3	1.4	10.1	2.2	3.3	0.8
Male	18.3	4.6	12.6	2.9	10.6	2.7	7.2	1.9	10.3	2.3	3.6	1.0
Female	14.8	3.5	12.6	2.8	10.3	1.6	7.4	0.9	9.9	2.1	3.1	0.3
Total	16.5	4.3	12.6	2.9	10.6	2.3	6.8	1.2	9.5	2.2	3.5	0.9

APPENDIX C

SBCL Item Loadings with Principal Factoring with Iteration and Varimax Rotation--6 Factors

	Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
2.	Tends to give up	.53054	.37180	.18510	-.06181	.24073	.01784
3.	Interrupts	.04695	.40403	.26777	.44948	-.07086	-.02692
5.	Fights over nothing	.16210	.16446	.48526	.15720	-.11865	.07328
7.	Alert in class	.57176	-.07280	.31515	.14015	.36040	.05365
12.	Volunteers to recite	.04921	.06609	.15769.	.10045	.36827	.00105
13.	Hits and pushes	.11310	.15087	.78661	.06341	-.00119	.04237
17.	Not considerate	.23421	.17243	.54817	.16751	-.12085	.12244
18.	Fails to carry out tasks	.84293	.15160	.15998	.14527	-.01880	-.07383
19.	Lacks ambition	.79505	.14036	.16733	.08516	.04621	-.05267
20.	Does things to anger others	.02499	.22840	.64130	.33061	.19972	.02154
21.	Argues when told not to do something	.19338	.61924	.26332	.24028	-.12655	.09555

APPENDIX C (cont'd)

Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
22. Does homework	.88023	.12864	.11125	.04889	.05566	.03932
23. Teases other children	.09142	.30947	.64387	.22262	.16341	-.00267
26. Easily upset by change	.09542	.51242	.12041	.04935	.13521	.22064
27. Is sure of self	.50129	.19879	-.01588	-.03478	.49244	.07463
28. Uses abusive language	-.01911	.25514	.47610	.33940	.16513	.29415
29. Has changeable moods	.11228	.58827	.13621	.26470	-.05052	-.09282
31. Doesn't respect others belongings	.23911	.23912	.55854	.07270	-.03767	.05867
34. Infuriated by discipline	.23901	.38398	.27119	.17732	.15754	.05194
35. Likes an audience	.23650	.11325	.11376	.77291	-.04810	-.02737
36. Finds it hard to study	.79478	.06065	.17679	.23438	.21259	.04123
38. Works well alone	.47883	.28839	.09356	.18909	.35226	-.01723
39. When angry refuses to speak	.17819	.47127	.18796	-.04523	.02173	-.03459

APPENDIX C (cont'd)

Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
43. Seems dull; Slow to catch on	.29187	.00287	.08519	.13984	.50539	.02474
44. Won't ask questions	.17447	-.06248	.11243	.00270	.46597	.04464
46. Never seems to be still	.24140	.13252	.38402	.47152	-.02830	.08896
47. Argues with teacher	.13634	.63798	.21963	.38611	-.03639	.01596
48. Able to concentrate	.61085	.07384	.02606	.15571	.36565	.33359
51. Tries to be center of attention	.05805	.36652	.19061	.72021	-.03282	-.01252
52. "Drags feet" when requested to do something	.72052	.18660	-.05023	.07954	.00542	.12441
54. Sulks	.19059	.64986	.09126	.30017	.01859	-.01427
57. Distractible	.67789	.12575	.16341	.17590	.24142	.16507
62. Stubborn	.28177	.46700	.38528	.35021	-.15002	-.03904
63. Never speaks up	.02071	-.22165	-.06074	-.07632	.43501	.09536
64. Interested in school work	.64642	.05580	.18032	.06949	-.01817	.35858

APPENDIX C (cont'd)

	Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
66.	Does things to attract attention	.13067	.42369	.38173	.45412	.21131	-.00676
69.	Popular	.39924	.00644	.23988	-.00609	.21604	.70871
70.	Does things normal for younger children	.08179	-.2929	-.08342	.58303	.30619	.26105
72.	Threatens other children	.09189	.34008	.50510	.42314	-.03593	.00697
75.	Prefers to play and be alone	.00109	-.03723	-.17228	.10096	.01769	.33048
76.	Finishes assignments	.79733	.04430	.09661	.21138	-.00047	.20295
77.	Gives dirty looks	.01543	.51233	.25475	.37909	.01162	.11920
78.	Deliberately interrupts	.32116	.25775	.12591	.43487	-.06495	.13176
79.	Slow to make friends	.32330	.09732	.15090	-.07304	.18431	.85365
86.	When angry, slams door	.01993	.42743	.18100	.05610	-.02399	-.11590
87.	"Dare-devil"	.10673	.13724	.23457	.53403	-.03242	-.03154

	Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
88.	Difficulty speaking when excited	-.16812	.26453	.02917	-.14147	.43738	.06415
92.	Disturbs others with boisterousness	.19813	.17058	.38884	.52427	.11160	-.03318
16.	Defeatist attitude	.20469	.7712	-.00682	-.11890	.45720	.08744

APPENDIX D

Simple Correlations Between Individual Items of the SBCL and the Initial Stanford Reading Achievement Test (SAT_1)

Item		Total[a]	Boys[b]	Girls[c]
1.	Is friendly	-.02	.14	.17
2.	Tends to give up if has something hard to finish	-.48	-.49	-.38
3.	Interrupts whoever is speaking	-.17	-.31	.09
4.	Penmanship at least one grade level below age expectation	-.23	-.15	-.29
5.	Starts fighting over nothing	- 25	- 28	- 10
6.	Is a helpful child	21	19	17
7.	Is alert in class	40		
8.	Poorly coordinated when doing things with hands, such as coloring	- 15	- 17	18
9.	Reading ability at least one grade level below age expectation	- 65	- 66	- 59
10.	Just stands around on the playground	00	01	01
11.	Acts up when adults not watching	- 14	- 11	- 13
12.	Volunteers to recite in class	13	12	23
13.	Hits and pushes other children	- 21	- 23	- 12
14.	Hands shake or is nervous when called on to recite	- 08	- 04	- 09
15.	Finds fault with what other children do	- 10	- 14	- 11
16.	Approaches a difficult task with an air of defeatism	- 26	- 31	- 04
17.	Is considerate of others	10	13	00
18.	Fails to carry out tasks	- 36	- 39	- 17
19.	Lacks ambition to do well in school	- 34	- 35	- 18
20.	Does things to get others angry	- 20	- 29	00
21.	Will put up an argument when told not to do something	- 25	- 27	- 15
22.	Does homework	38	39	21
23.	Teases other children	- 29	- 34	- 23
24.	Is afraid of making mistakes	- 04	02	- 07
25.	Is bossy with other children	- 05	- 02	- 16
26.	Is easily upset by changes	- 16	- 10	- 10
27.	Is sure of self	40	41	32
28.	Uses abusive language toward other children	- 25	- 24	04
29.	Has changeable moods	- 22	- 34	04
30.	Gives in when another child insists on doing something another way	03	10	- 17
31.	Does not respect other people's belongings	- 24	- 25	- 13

APPENDIX D (cont'd)

Item		Total[a]	Boys[b]	Girls[c]
32.	Does not forget things which anger her/him	-12	-17	-08
33.	Seems to be off in own world	-16	-16	-09
34.	Is infuriated by any form of discipline	-20	-33	-01
35.	Likes an audience all the time	-31	-36	-01
36.	Finds it hard to study	-42	-45	-24
37.	Has to have everything own way	-03	-13	11
38.	Works well alone	27	28	19
39.	When angry, will refuse to speak to anyone	-17	-10	-32
40.	School performance is far below capabilities	-35	-36	-18
41.	Has no friends	10	14	06
42.	Behind at least one school grade due to academic difficulties	-54	-54	-49
43.	Seems dull; slow to catch on	-42	-44	-31
44.	Will not ask questions even when unsure as to how to do the work	-13	-23	05
45.	Fights back if another child has been asking for it	-05	-05	03
46.	Never seems to be still for a moment	-26	-26	-17
47.	Argues with me	-22	-29	-10
48.	Is able to concentrate on things	31	29	24
49.	Boasts of own toughness	-19	-15	-17
50.	Seems to think of self as worthless	-04	05	-19
51.	Tries to be the center of attention	-22	-32	05
52.	"Drags feet" when requested to do something	-23	-18	-25
53.	Accepts adult suggestions	10	06	12
54.	Sulks when things go wrong	-29	-34	-14
55.	Becomes frightened easily	-10	-09	-08
56.	Resents even the most gentle of criticism of work	-06	-11	12
57.	Distractible; can't concentrate	-37	-33	-37
58.	Able to see the bright side of things	05	01	07
59.	Fights with smaller children	-04	-04	-01
60.	Spelling performance at least one grade level below age expectation	-53	-57	-38
61.	Fearful of being hurt at play	07	-01	07
62.	Is stubborn	-24	-30	-08
63.	Never speaks up even when there is cause to be angry	-10	-28	12
64.	Is interested in schoolwork	38	42	21

APPENDIX D (cont'd)

Item		Total[a]	Boys[b]	Girls[c]
65.	Tries to get other children into trouble	04	05	00
66.	Does things just to attract attention	-30	-30	-23
67.	Never fights back, even if someone hits or pushes first	01	02	-04
68.	Prefers to be around adults rather than play with children	07	13	-04
69.	Is popular with his classmates	28	28	25
70.	Does things which are normal for children much younger	-30	-32	-16
71.	Never sticks up for self when picked on by other children	08	22	-11
72.	Threatens to hurt other children when angry	-23	-25	-06
73.	Average or above IQ	47	41	54
74.	Does not take orders when other children are in charge	-14	-20	01
75.	Prefers to be alone and play alone	15	28	-12
76.	Finishes classroom assignments	30	30	17
77.	Gives other children dirty looks	-16	-22	-09
78.	Deliberately interrupts what is going on by asking silly questions	-31	-35	01
79.	Slow in making friends	-26	-20	-33
80.	Seems as happy as most children	17	15	10
81.	Finds fault with instructions given by adults	-04	-04	-01
82.	Seems unconcerned when misbehaving	-20	-20	-17
83.	Cries easily	09	12	05
84.	Is afraid of strange adults	04	08	-01
85.	Is self-confident	17	13	20
86.	When angry, will do things like slamming the door or banging the desk	-18	-09	-34
87.	Acts in a "dare-devil", fearless manner	-21	-26	12
88.	Has difficulty speaking clearly when excited or upset	-12	-26	12
89.	Has a "chip on shoulder"	-07	-14	11
90.	Becomes embarrassed easily	-13	-15	-17
91.	Bright, but doesn't apply self	-16	-14	-06
92.	Disturbs other children with boisterous behavior	-18	-22	01
93.	Behind at least two school grades due to academic difficulties	-30	-29	-21
94.	Arithmetic skill at least one grade level below age expectation	-42	-41	-38

APPENDIX D (cont'd)

Item	Total[a]	Boys[b]	Girls[c]
95. Much anxiety--afraid of such things as storms, school, death, injury, war	-05	-03	-02
96. Frequent headaches, stomach aches or other non-specific physical complaints	02	11	-13

(a) r .23 significant at .01 level for total group
(b) r .28 significant at .01 level for
(c) r .31 significant at .01 level for

APPENDIX E

Simple Correlations Between Individual Items of the SBCL and the Stanford Reading Achievement Test (SAT_1, SAT_2, SAT_3, SAT_4) (Total Group)

Item	SAT_1	SAT_2	SAT_3	SAT_4
1. Is friendly	-02	12	-03	-04
2. Tends to give up if has something hard to finish	-48	-44	-45	-48
3. Interrups whoever is speaking	-17	-12	-17	-13
4. Penmanship (handwriting) at least one grade level below age expectation	-23	-27	-14	-22
5. Starts fighting over nothing	-25	-25	-14	-22
6. Is a helpful child	21	27	15	16
7. Is alert in class	40	41	36	28
8. Poorly coordinated when doing things with hands, such as coloring or pencil work	-15	-19	-03	-02
9. Reading ability at least one grade level below age expectation	-65	-63	-57	-60
10. Just stands around on the playground	00	-04	-01	03
11. Acts up when adults not watching	-14	-15	-20	-33
12. Volunteers to recite in class	29	25	18	24
13. Hits and pushes other children	-21	-27	-17	-27
14. Hands shake or is nervous when called on to recite	-08	-03	04	-08
15. Finds fault with what other children do	-10	-15	-15	-23
16. Approaches a difficult task with an air of defeatism	-26	-20	-19	-19
17. Is considerate of others	10	19	10	15
18. Fails to carry out tasks (homework assignments, seat work, etc.)	-36	-36	-32	-28
19. Lacks ambition to do well in school	-34	-34	-21	-24
20. Does things to get others angry	-20	-27	-21	-26
21. Will put up an argument	-25	-21	-24	-36
22. Does homework	38	42	32	24
23. Teases other children	-29	-31	-27	-33
24. Is afraid of making mistakes	-04	-03	01	-02
25. Is bossy with other children	-05	-07	-10	-14
26. Is easily upset by changes	-16	-08	-10	-21

APPENDIX E (cont'd)

Item		SAT_1	SAT_2	SAT_3	SAT_4
27.	Is sure of self	40	38	26	34
28.	Uses abusive language toward other children	-25	-19	-24	-23
29.	Has changeable moods	-22	-25	-28	-25
30.	Gives in when another child insists on doing something another way	03	02	08	07
31.	Does not respect other people's belongings	-24	-21	-23	-23
32.	Does not forget things which anger her/him	-21	-04	-11	-11
33.	Seems to be off in own world	-16	-30	-18	-07
34.	Is infuriated by any form of discipline	-20	-16	-19	-25
35.	Likes an audience all the time	-31	-32	-24	-25
36.	Finds it hard to study	-42	-44	-32	-35
37.	Has to have everything own way	-03	-10	-14	-01
38.	Works well alone	27	28	25	19
39.	When angry, will refuse to speak to anyone	-17	-21	-23	-27
40.	School performance is far below capabilities	-35	-39	-18	-11
41.	Has no friends	10	12	11	15
42.	Behind at least one school grade due to academic difficulties	-54	-54	-42	-33
43.	Seems dull; slow to catch on	-42	-44	-36	-34
44.	Will not ask questions even when unsure as to how to do the work	-23	-23	-16	-19
45.	Fights back if another child has been asking for it	-05	-03	-03	-07
46.	Never seems to be still for a moment	-26	-20	-27	-28
47.	Argues with me	-22	-22	-24	-30
48.	Is able to concentrate on things	31	35	24	21
49.	Boasts of own toughness	-19	-11	-18	-26
50.	Seems to think of self as worthless	-04	-05	06	05
51.	Tries to be the center of attention	-22	-21	-22	-23

APPENDIX E (cont'd)

Item		SAT_1	SAT_2	SAT_3	SAT_4
52.	"Drags feet" when requested to do something	-23	-27	-13	-15
53.	Accepts adult suggestions	10	19	11	11
54.	Sulks when things go wrong	-29	-31	-28	-38
55.	Becomes frightened easily	-10	-09	02	-07
56.	Resents even the most gentle criticism of work	-06	-06	-06	-10
57.	Distractible; can't concentrate	-37	-40	-27	-21
58.	Able to see the bright side of things	05	02	05	14
59.	Fights with smaller children	-04	-04	-03	-15
60.	Spelling performance at least one grade level below age expectation	-53	-56	-39	-42
61.	Fearful of being hurt at play	07	18	18	18
62.	Is stubborn	-27	-21	-18	-23
63.	Never speaks up even when there is cause to be angry	-10	-15	-08	-06
64.	Is interested in schoolwork	38	38	29	30
65.	Tries to get other children in trouble	04	-12	-04	-02
66.	Does things just to attract attention	-30	-28	-33	-40
67.	Never fights back, even if someone hits or pushes first	-01	-01	-05	-06
68.	Prefers to be around adults, rather than play with children	-17	-23	-16	-11
69.	Is popular with his classmates	28	38	25	24
70.	Does things which are normal for children much younger	-30	-30	-17	-24
71.	Never sticks up for self	08	05	09	06
72.	Threatens to hurt other children when angry	-23	-28	-24	-32
73.	Average or above IQ	-47	-46	-46	-50
74.	Does not take orders when other children are in charge	-14	-16	-17	-23
75.	Prefers to be alone and play alone	25	16	08	18
76.	Finishes classroom assignments	30	28	18	25
77.	Gives other children dirty looks	-16	-18	-21	-32
78.	Deliberately interrupts what is going on by asking silly questions	-31	-28	-26	-30

APPENDIX E (cont'd)

Item		SAT_1	SAT_2	SAT_3	SAT_4
79.	Slow in making friends	-26	-35	-22	-18
80.	Seems as happy as most children	17	23	08	05
81.	Finds fault with instructions given by adults	-04	-09	-15	-26
82.	Seems unconcerned when misbehaving	-20	-23	-20	-29
83.	Cries easily	09	14	10	01
84.	Is afraid of strange adults	04	01	00	06
85.	Is self-confident	17	18	12	06
86.	When angry, will do things like slamming the door	-18	-17	-14	-29
87.	Acts in a "dare-devil", fearless manner	-21	-23	-14	-19
88.	Has difficulty speaking clearly when excited or upset	-12	-13	01	-12
89.	Has a "chip on shoulder"	-07	-16	-09	-17
90.	Becomes embarrassed easily	-13	-11	-07	-03
91.	Bright, but doesn't apply self (underachiever)	-16	-19	-11	-17
92.	Disturbs other children with boisterous behavior	-18	-16	-16	-23
93.	Behind at least two school grades due to academic difficulties	-30	-28	-07	-18
94.	Arithmetic skill at least one grade level below age expectations	-42	-46	-30	-35
95.	Much anxiety--afraid of such things as storms, school, death, injury, was (considered phobic)	-05	-10	07	00
96.	Frequent headaches, stomach aches or other non-specific physical complaints	02	-02	12	-16

* r .23 significant at .01 level

APPENDIX F

Simple Correlations Between Individual Items of the SBCL and the Stanford Achievement Test (SAT_1, SAT_2, SAT_3, SAT_4) for Boys and Girls

Item	SAT_1		SAT_2		SAT_3		SAT_4	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
1. Is friendly	14	17	17	24	16	14	22	17
2. Tends to give up if has something hard to finish	-49	-38	-49	-26	-47	-34	-46	-50
3. Interrupts whoever is speaking	-31	09	-29	16	-33	05	-29	08
4. Penmanship at least one grade level below age expectation	-15	-29	-22	-29	-01	-32	-14	-36
5. Starts fighting over nothing	-28	-10	-24	-22	-16	-09	-33	-00
6. Is a helpful child	19	17	26	24	12	14	14	17
7. Is alert in class	39	32	40	37	31	39	34	14
8. Poorly coordinated when doing things with hands, such as coloring	-17	18	-28	32	-06	35	-07	28
9. Reading ability at least one grade level below age expectation	-66	-58	-66	-53	-56	-53	-56	-66
10. Just stands around on the playground	00	01	00	-08	08	-16	12	-12
11. Acts up when adults not watching	-11	-13	-07	-25	-18	-18	-32	-31
12. Volunteers to recite in class	12	23	12	32	-05	27	05	31
13. Hits and pushes other children	-23	-12	-29	-20	-23	-02	-40	-05
14. Hands shake or is nervous when called on to recite	-04	-09	-01	04	16	-15	-07	-08
15. Finds fault with what other children do	-14	-11	-16	-20	-20	-17	-27	-22
16. Approaches a difficult task with an air of defeatism	-31	-04	-29	06	-21	-07	-23	-05
17. Is considerate of others	13	00	19	12	14	-02	19	05
18. Fails to carry out tasks	-39	-17	-39	-20	-30	-26	-25	-28
19. Lacks ambition to do well in school	-35	-18	-33	-29	-14	-25	-16	-35
20. Does things to get others angry	-29	00	-33	-16	-30	-03	-37	-09

APPENDIX F (cont'd)

Item		SAT_1		SAT_2		SAT_3		SAT_4	
		Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
21.	Will put up an argument when told not to do something	27	15	21	18	23	23	34	36
22.	Does homework	39	-21	-45	-26	-29	-29	-18	-33
23.	Teases other children	-34	-16	-34	-23	-38	-09	-41	-21
24.	Is afraid of making mistakes	02	-15	08	-07	13	-16	-08	04
25.	Is bossy with other children	-02	-13	-03	-16	-11	-10	-03	-27
26.	Is easily upset by changes	-10	-23	-04	-10	00	-21	-15	-28
27.	Is sure of self	41	31	39	32	20	29	35	29
28.	Uses abusive language	-24	-19	-25	04	-25	-15	-29	-09
29.	Has changeable moods	-34	06	-38	04	-41	-04	-41	-02
30.	Gives in when another child insists on doing something another way	10	-18	09	-17	21	-16	08	01
31.	Does not respect other people's belongings	-25	-13	-25	-09	-29	-07	-29	-10
32.	Does not forget things which anger her/him	-17	-08	-13	08	-21	-05	-23	01
33.	Seems to be off in own world	-16	-09	-36	-12	-16	-13	-10	03
34.	Is infuriated by any form of discipline	-33	01		03	-32	-03	-34	-13
35.	Likes an audience all the time	-36	-01	-35	-14	-24	-14	-26	-22
36.	Finds it hard to study	-45	-24	-50	-25	-28	-30	-36	-31
37.	Has to have everything own way	-13	11	-10	-13	-25	-01	-18	00
38.	Works well alone	28	19	32	15	16	32	10	26
39.	When angry, will refuse to speak to anyone	-10	-32	-27	-19	-23	-26	-17	-39
40.	School performance is far below capabilities	-36	-18	-41	-27	-14	-17	-13	02
41.	Has no friends	-14	-02	14	-08	16	05	22	08
42.	Behind at least one school grade due to academic difficulties	-54	-49	-54	-49	-40	-39	-32	-30
43.	Seems dull; slow to catch on	-44	-31	-50	-24	-37	-27	-35	-30
44.	Will not ask questions even when unsure as to how to do the work	-23	-5	-20	02	-10	00	-18	02

APPENDIX F (cont'd)

Item		SAT_1		SAT_2		SAT_3		SAT_4	
		Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
45.	Fights back if another child has been asking for it	-05	03	-02	18	-09	12	-07	-02
46.	Never seems to be still for a moment	-26	-17	-19	-17	-26	-22	-30	-22
47.	Argues with me	-29	-10	-28	-11	-29	-18	-31	-28
48.	Is able to concentrate on things	29	24	39	25	14	35	23	13
49.	Boasts of own toughness	-15	-17	-13	12	-16	-16	-31	-13
50.	Seems to think of self as worthless	05	-19	04	-26	20	-22	00	28
51.	Tries to be the center of attention	-32	05	-30	04	-28	-07	-33	-07
52.	"Drags feet" when requested to do something	-18	-25	-22	-32	-02	-30	-08	-47
53.	Accepts adult suggestions	06	12	09	35	06	34	-07	37
54.	Sulks when things go wrong	-34	-34	-34	-23	-36	-12	-46	-27
55.	Becomes frightened easily	-09	05	-08	04	07	-07	-07	-01
56.	Resents even the most gentle criticism of work	-11	12	-06	-01	-08	03	-14	04
57.	Distractible; can't concentrate	-33	-37	-36	-41	-17	-35	-14	-26
58.	Able to see the bright side of things	01	-07	-02	-10	02	-09	16	-12
59.	Fights with smaller children	-04	-01	-09	12	-04	05	-33	04
60.	Spelling performance at least one grade level below age expectation	-57	-38	-63	-37	-41	-26	-40	-42
61.	Fearful of being hurt at play	-01	07	08	25	13	18	08	21
62.	Is stubborn	-30	-08	-29	-06	-25	-02	-29	-15
63.	Never speaks up even when there is cause to be angry	-28	12	-32	03	-16	00	-18	05
64.	Is interested in schoolwork	42	21	39	32	27	27	26	36
65.	Tries to get other children in trouble	05	00	-09	-18	01	-11	02	-09
66.	Does things which are normal for children much younger	-30	-23	-29	-21	-34	-28	-42	-36
67.	Never fights back, even if someone hits or pushes first	02	-04	-03	-02	-05	-09	-07	-09
68.	Prefers to be around adults rather than play with children	13	-04	-01	-09	11	-01	09	10

APPENDIX F (cont'd)

Item		SAT_1		SAT_2		SAT_3		SAT_4	
		Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
69.	Is popular with his classmates	28	25	37	39	17	34	25	19
70.	Does things which are normal for children much younger	-33	-16	-28	-32	-27	-32	-36	-24
71.	Never sticks up for self when picked on by other children	22	-11	15	-08	22	-08	21	-10
72.	Threatens to hurt other children when angry	-25	-06	-32	-09	-32	03	-47	-02
73.	Average or above IQ (Intelligence Quotient)	41	54	46	41	41	51	53	47
74.	Does not take orders when other children are in charge	-20	01	-18	-09	-21	-06	-27	-17
75.	Prefers to be alone and play alone	28	-12	18	-17	22	-15	24	14
76.	Finishes classroom assignments	30	17	28	20	08	26	20	28
77.	Gives other children dirty looks	-21	-09	-22	-16	-26	-17	-36	-29
78.	Deliberately interrupts what is going on by asking silly questions	-35	01	-32	-02	-32	05	-38	03
79.	Slow in making friends	-20	-33	-33	-39	-14	-33	-16	-18
80.	Seems as happy as most children	15	10	20	23	01	14	12	28
81.	Finds fault with instructions given by adults	-04	-01	-09	-06	-10	-22	-30	-19
82.	Seems unconcerned when misbehaving	-20	-17	-20	-26	-14	-27	-28	-29
83.	Cries easily	12	05	19	07	16	02	04	-02
84.	Is afraid of strange adults	08	-01	-01	03	04	-03	-12	03
85.	Is self-confident	13	20	17	16	03	22	04	05
86.	When angry, will do things like slamming the door or banging the desk	-09	-34	-11	-26	-06	-24	-21	-38
87.	Acts in a "dare-devil", fearless manner	-26	19	-27	10	-20	17	-30	11
88.	Has difficulty speaking clearly when excited or upset	-04	-28	-04	-29	08	-09	-05	-19
89.	Has a "chip on shoulder"	-14	11	-23	00	-21	11	-33	13

APPENDIX F (cont'd)

Item		SAT_1		SAT_2		SAT_3		SAT_4	
		Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
90.	Becomes embarrassed easily	-15	-17	-08	-20	-05	-16	-15	07
91.	Bright, but doesn't apply self (underachiever)	-14	-06	-16	-17	-04	-13	-11	-22
92.	Disturbs other children with boisterous behavior	-22	01	-20	01	-19	-04	-30	-06
93.	Behind at least two school grades due to academic difficulties	-29	-21	-27	-20	-01	-10	-17	-17
94.	Arithmetic skill at least one grade level below age expectation	-41	-38	-49	-35	-21	-38	-26	-46
95.	Much anxiety--afraid of such things as storms, school, death, injury, was (considerated phobic)	-03	-05	-09	-08	12	-03	01	00
96.	Frequent headaches, stomach aches or other non-specific physical complaints	11	-13	03	-09	21	01	-13	-20

APPENDIX G

Relative Frequency of Student Behavior Attributes Based on Teachers' Ratings From the School Behavior Checklist

Item		No	Yes
1.	Is friendly	2%	98%
2.	Tends to give up if something hard to finish	65	35
3.	Interrupts whoever is speaking	73	27
4.	Penmanship at least one grade level below age expectation	78	22
5.	Starts fighting over nothing	88	12
6.	Is a helpful child	7	93
7.	Is alert in class	23	77
8.	Poorly coordinated when doing things with hands, such as coloring or pencil work	88	12
9.	Reading ability at least one grade level below age expectation	68	32
10.	Just stands around on the playground	96	4
11.	Acts up when adults not watching	70	30
12.	Volunteers to recite in class	19	81
13.	Hits and pushes other children	78	22
14.	Hands shake or is nervous when called on to recite	89	11
15.	Finds fault with what other children do	73	27
16.	Approaches a difficult task with an air of defeatism	85	15
17.	Is considerate of others	19	81
18.	Fails to carry out tasks	74	26
19.	Lacks ambition to do well in school	77	23
20.	Does things to get others angry	77	23
21.	Will put up an argument when told not to do something	72	28
22.	Does homework	24	76
23.	Teases other children	72	28
24.	Is afraid of making mistakes	79	21
25.	Is bossy with other children	82	18
26.	Is easily upset by changes	81	19
27.	Is sure of self	38	62
28.	Uses abusive language toward other children	87	13
29.	Has changeable moods	57	43
30.	Gives in when another child insists on doing something another way	55	45
31.	Does not respect other people's belongings	92	8
32.	Does not forget things which anger her/him	91	9

APPENDIX G (cont'd)

Item		No	Yes
33.	Seems to be off in own world	83	17
34.	Is infuriated by any form of discipline	87	13
35.	Likes an audience all the time	86	14
36.	Finds it hard to study	74	26
37.	Has to have everything own way	87	13
38.	Works well alone	36	64
39.	When angry, will refuse to speak to anyone	87	13
40.	School performance is far below capabilities	80	20
41.	Has no friends	99	1
42.	Behind at least one school grade due to academic difficulties	75	25
43.	Seems dull; slow to catch on	88	12
44.	Will not ask questions even when unsure as to how to do the work	91	9
45.	Fights back if another child has been asking for it	30	70
46.	Never seems to be still for a moment	79	21
47.	Argues with me	70	30
48.	Is able to concentrate on things	22	78
49.	Boasts of own toughness	88	12
50.	Seems to think of self as worthless	92	8
51.	Tries to be the center of attention	78	22
52.	"Drags feet" when requested to do something	74	26
53.	Accepts adult suggestions	19	81
54.	Sulks when things go wrong	65	35
55.	Becomes frightened easily	96	4
56.	Resents even the most gentle criticism of work	93	7
57.	Distractible; can't concentrate	71	29
58.	Able to see the bright side of things	6	94
59.	Fights with smaller children	96	4
60.	Spelling performance at least one grade level below age expectation	68	32
61.	Fearful of being hurt at play	91	9
62.	Is stubborn	62	38
63.	Never speaks up even when there is cause to be angry	94	6
64.	Is interested in schoolwork	15	85
65.	Tries to get other children into trouble	92	8
66.	Does things just to attract attention	74	26
67.	Never fights back, even if someone hits or pushes firts	93	7
68.	Prefers to be around adults, rather than play with children	94	6

APPENDIX G (cont'd)

Item		No	Yes
69.	Is popular with his classmates	18	82
70.	Does things which are normal for children much younger	90	10
71.	Never sticks up for self when picked on by other children	94	6
72.	Threatens to hurt other children when angry	76	24
73.	Average or above IQ	28	72
74.	Does not take orders when other children are in charge	83	17
75.	Prefers to be alone and play alone	91	9
76.	Finishes classroom assignments	26	74
77.	Gives other children dirty looks	72	28
78.	Deliberately interrupts what is going on by asking silly questions	88	12
79.	Slow in making friends	82	18
80.	Seems as happy as most children	12	88
81.	Finds fault with instructions given by adults	87	13
82.	Seems unconcerned when misbehaving	78	22
83.	Cries easily	91	9
84.	Is afraid of strange adults	94	6
85.	Is self-confident	22	78
86.	When angry, will do things like slamming the door or banging the desk	84	16
87.	Acts in a "dare-devil", fearless manner	86	14
88.	Has difficulty speaking clearly when excited or upset	80	20
89.	Has a "chip on shoulder"	90	10
90.	Becomes embarrassed easily	72	28
91.	Bright, but doesn't apply self	78	22
92.	Disturbs other children with boisterous behavior	88	12
93.	Behind at least two school grades due to academic difficulties	89	11
94.	Arithmetic skill at least one grade level below age expectation	69	31
95.	Much anxiety--afraid of such things as storms, school, death, injury, war	97	3
96.	Frequent headaches, stomach aches or other non-specific physical complaints	93	7

APPENDIX H

The School Behavior Checklist - Form A2

School Behavior Checklist

by Lovick C. Miller, Ph.D.

Form A2

Ages 7-13

Published by

WPS WESTERN PSYCHOLOGICAL SERVICES
PUBLISHERS AND DISTRIBUTORS
12031 WILSHIRE BOULEVARD
LOS ANGELES, CALIFORNIA 90025
A DIVISION OF MANSON WESTERN CORPORATION

DIRECTIONS

The statements on this checklist are often used by teachers to describe children's behavior. Read each statement and decide whether it describes the child being rated. If it does, blacken the space for that item under "T" on the separate answer sheet; if not, blacken the space under "F" on the answer sheet. It is important to mark EACH statement. If you are in doubt, blacken the answer that is most characteristic of the child.

STATEMENTS

1. Is friendly
2. Tends to give up if has something hard to finish
3. Interrupts whoever is speaking
4. Penmanship (handwriting) at least one grade level below age expectation
5. Starts fighting over nothing
6. Is a helpful child
7. Is alert in class
8. Poorly coordinated when using hands, such as in coloring or pencil work
9. Reading ability at least one grade level below age expectation
10. Just stands around on the playground
11. Acts up when adults not watching
12. Volunteers to recite in class
13. Hits and pushes other children
14. Hands shake or is nervous when called on to recite
15. Finds fault with what other children do
16. Approaches a difficult task with an air of defeatism
17. Is considerate of others
18. Fails to carry out tasks (homework assignments, seat work, etc.)
19. Lacks ambition to do well in school
20. Does things to get others angry
21. Will put up an argument when told not to do something
22. Does homework
23. Teases other children
24. Is afraid of making mistakes
25. Is bossy with other children
26. Is easily upset by changes
27. Is sure of self
28. Uses abusive language toward other children
29. Has changeable moods
30. Gives in when another child insists on doing something another way
31. Does not respect other people's belongings
32. Does not forget things which anger her/him
33. Seems to be off in own world
34. Is infuriated by any form of discipline
35. Likes an audience all the time
36. Finds it hard to study
37. Has to have everything own way
38. Works well alone
39. When angry, will refuse to speak to anyone
40. School performance is far below capabilities

W-144A

APPENDIX H (cont'd)

41. Has no friends
42. Behind at least one school grade due to academic difficulties
43. Seems dull; slow to catch on
44. Will not ask questions even when unsure as to how to do the work
45. Fights back if another child has been asking for it
46. Never seems to be still for a moment
47. Argues with me
48. Is able to concentrate on things
49. Boasts of own toughness
50. Seems to think of self as worthless
51. Tries to be the center of attention
52. "Drags feet" when requested to do something
53. Accepts suggestions of adults
54. Sulks when things go wrong
55. Becomes frightened easily
56. Resents even the most gentle criticism of work
57. Distractible; can't concentrate
58. Able to see the bright side of things
59. Fights with smaller children
60. Spelling performance at least one grade level below age expectation
61. Fearful of being hurt at play
62. Is stubborn
63. Never speaks up even when there is cause to be angry
64. Is interested in schoolwork
65. Tries to get other children into trouble
66. Does things just to attract attention
67. Never fights back, even if someone hits or pushes first
68. Prefers to be around adults, rather than play with childern
69. Is popular with classmates
70. Does things that are normal for children much younger
71. Never sticks up for self when picked on by other children
72. When angry, threatens to hurt other children
73. IQ (Intelligence Quotient) average or above
74. Does not take orders when other children are in charge
75. Prefers to be alone and play alone
76. Finishes classroom assignments
77. Gives other children dirty looks
78. Deliberately interrupts what is going on by asking silly questions
79. Slow in making friends
80. Seems as happy as most children
81. Finds fault with instructions given by adults
82. Seems unconcerned when misbehaving
83. Cries easily
84. Is afraid of strange adults
85. Is self-confident
86. When angry, will do things like slamming the door or banging the desk
87. Acts in a "dare-devil," fearless manner
88. Has difficulty speaking clearly when excited or upset
89. Has a "chip on shoulder"
90. Becomes embarrassed easily
91. Bright, but doesn't apply self (underachiever)
92. Disturbs other children with boisterous behavior
93. Behind at least two school grades due to academic difficulties
94. Arithmetic skill at least one grade level below age expectation
95. Unusually afraid of such things as storms, school, death, injury (phobic)
96. Frequent headaches, stomach aches or other non-specific physical complaints

APPENDIX I

Permission Forms for Release of Information

MICHIGAN STATE UNIVERSITY EAST LANSING · MICHIGAN 48823

COLLEGE OF EDUCATION · ERICKSON HALL

July 15, 1981

Dear Parent(s):

In the fall of 1977 your child's classroom at Maple Grove Elementary School participated in a study conducted by Michigan State University and the Lansing School District. The purpose of this study was to develop a better understanding of how children learn to read. Parental consent forms were sent to your home at that time, and you allowed us to obtain information regarding your child's progress in reading.

We would like to follow-up on our earlier findings. To do this, we need your consent to obtain information regarding your child's current progress in reading as well as his/her growth in reading over the past four years. This information has already been obtained by the Lansing School District.

In making use of this information anonymity will be protected: your child will not be identified by name in any aspect of the study. If you approve of our use of this information, please sign this letter and return it by July 30, 1981 in the stamped self-addressed envelope.

Sincerely,

Robert L. Jarvis

Research Assistant

Richard S. Prawat

Assosiate Professor
Senior Research Assistant

_____I approve of releasing my child's reading achievement information.

Parent(s):_____

Child:_____

APPENDIX I (cont'd)

MICHIGAN STATE UNIVERSITY EAST LANSING · MICHIGAN 48823

COLLEGE OF EDUCATION · ERICKSON HALL

July 29, 1981

Dear Parent(s):

Within the past two weeks you received a letter requesting release of your child's reading achievement scores from the Lansing School District for purposes of a follow-up study in conjunction with Michigan State University. Because this is a follow-up study, much of its practical importance depends upon gaining current reading achievement information on as many of the students in the original 1977 study as possible.

If you have already returned your permission form, releasing your child's reading scores, we sincerely appreciate your cooperation and concern in helping us better understand those factors which contribute to reading success or failure among students in Lansing. If you have not as yet returned this form, we would again like to encourage your participation in this meaningful study by completing and returning the enclosed form.

We stress that your child will not be identified by name in any aspect of the study and the confidentiality of your child's name and test scores will be strictly guarded.

We again thank you for allowing your child's participation in our original 1977 study and encourage your support in our current efforts.

Sincerely,

Robert L. Jarvis
Research Assistant

Richard S. Prawat
Associate Professor

____ I approve of releasing my child's reading achievement information

Parent(s): _____

Child: _____