THE PREDICTIVE VALIDITY OF THREE PRESCHOOL SCREENING BATTERIES

by

Sandra F. Dill

Dissertation submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Counseling/Student Personnel Services

(School Psychology)

APPROVED:

Thomas Hohenshil, Co-Chair

Kusum Singh, Co-Chair

Libby Hoffman

Jack Finney

Cynthia Howard

December, 1995

Blacksburg, Virginia

Key Words: Preschool Screening, Predictive Validity, Reading, Early Determinants

THE PREDICTIVE VALIDITY
OF THREE PRESCHOOL SCREENING BATTERIES

by

Sandra Fortenberry Dill

Co-chairs: Thomas Hohenshil and Kusum Singh

Counseling/Student Personnel Services (School Psychology)

(ABSTRACT)

This study addressed the prevention of reading problems by examining the predictive validity of the Developmental Indicators of Assessment and Learning-Revised (DIAL-R), the Gesell School Readiness Test, and the Preschool Screening Battery. The effects of entry-level developmental abilities on reading performance were also studied.

Subjects were sixth grade students from five school systems in North Carolina. School system selection was based on the type of screening test used in 1988, and availability of data. Data were collected through student record reviews. Predictive discriminant analyses were computed to assess the accuracy of the screening batteries in predicting reading group membership at the third and fifth grade levels. Classification accuracy was determined by comparing hit ratios from the classification matrices with proportional chance criterion. Standard normal statistics and Press's Q statistics were computed to determine if the hit rates were better than those expected by chance. Early determinants were identified through interpretation of the discriminant

functions. Significant differences between correctly and incorrectly classified students were identified through t-tests comparing means scores of both groups.

Several implications were derived from the study. First, the magnitude of the discriminatory power of the DIAL-R and Preschool Screening Battery to predict reading group membership at the third and fifth grades was sufficient to recommend their continued use in preschool screening. The Gesell School Readiness Test was better than chance on only three out of twelve criteria, however, and is not recommended as a predictor of reading achievement.

Discriminant analyses resulted in the identification of three early determinants, conceptual development, biological development (fine motor and perceptual skills), and language comprehension. Expressive language emerged as the most important language subskill studied and auditory memory was the most important perceptual skill.

The multiple cause theory of reading acquisition was supported by these results. Visual motor skills were more influential in the early stages of word analysis, and language and conceptual development were more influential in predicting comprehension skills.

Intra-test variation in predictor variables appeared to be the most important contributor to inaccurate classification. Other factors were individual subject characteristics such as behavior problems or chronic illness.

ACKNOWLEDGEMENTS

The first person I wish to credit with influencing me in this academic pursuit is my father, William Fortenberry, who doggedly pursued his own college education over an eighteen year period through the depression years, a world war, and the birth of four children. I was profoundly affected by his respect for the value of a college education.

I wish to acknowledge my two sons, Scott and Kevin, who were with me as five and seven year olds when I received my B. S. degree, and who were also there for me when I pursued my M. S degree, and who accepted my busy schedule and hours spent studying without complaint. They were also there for me as adults when I decided to resign from my job and to pursue this degree as a mid-life endeavor. I thank them for their support and encouragement throughout all the years and for being the warm and loving individuals they have become.

I have been blessed to have had three excellent advisers over the four year period I have been a student at Virginia Tech. I wish to acknowledge Dr. Thomas Hohenshil to whom I will always be indebted for graciously agreeing to become my adviser in mid-stream and for guiding me through the prospectus exam and through the completion of the first three chapters of my dissertation. I thank Dr. Tim Keith for agreeing to be my adviser upon admission to the program and thus allowing me to begin this life-long pursuit. And finally, I am indebted to Dr. Kusum Singh, who became a co-chair, and spent many hours with me over the summer and fall advising me and guiding me through the data analysis and revisions of the final chapters.

Many thanks to my remaining committee members: Dr. Libby Hoffman, Dr. Jack Finney, and Dr. Cynthia Howard who supported me, sometimes through the trials of telephone communication, through the entire project. I acknowledge the support of the North Carolina Department of Instruction for allowing Dr. Howard to participate as a committee member.

Finally, I thank my friend, Chuck, for his patience and support over the last two years.

TABLE OF CONTENTS

LIST OF TABLES IN TEXT

LIST OF TABLES IN APPENDICES

CHAPTER I

Increased interest in early childhood education, early intervention, and early screening programs may be inferred from the expanded research and the growth of programs in these areas during the last three decades. Research in compensatory education, with programs such as Chapter I and Headstart, increased as rapidly as funding increased. At the same time, Public Law 94-142, the Education for the Handicapped Act, passed in 1974, mandated a comprehensive program of screening and identification of all handicapped children as well as a free and appropriate education for them. Compensatory programs funded during the 60's and 70's focused on improving educational and economic opportunities for low-income children (Stallings & Stepek, 1986).

Public Law 99-457, the Education of the Handicapped Act Amendments, passed in 1986 mandated services for handicapped children from birth to age 5 (Gowen, 1989), and extended the rights guaranteed by PL 94-142 to the preschool child. Passage of this law, again, rekindled interest in early screening, identification, and intervention.

During the same time period, the North Carolina State Legislature expressed concern regarding the developmental appropriateness of kindergarten programs in North Carolina by commissioning a study on kindergarten practices in the state (Bryant, Clifford & Peisner, 1989). Researchers commissioned to conduct the study found that 8.6% of the children in 53 kindergarten classes observed in the Spring of 1988 were retained at the end of kindergarten. Assuming an 8.6% statewide kindergarten retention rate, 7,310 children out of the statewide kindergarten population of 85,000 would have been retained during 1988.

The failure of many children to attain the readiness levels considered prerequisite for promotion into first grade highlighted the critical nature of the initial stages of skill acquisition. Prevention of the negative emotional consequences of learning failure seems imperative at this stage in a child's life. The effects of feelings such as defeat, anger, and frustration are easier to prevent than to ameliorate (Hagin, 1984; Stanovich, Cunningham & Feeman, 1984).

In another move, the North Carolina General Assembly passed the Basic Education Plan (BEP), a statewide initiative to equalize education for all children. In this plan, funds were appropriated to local school systems to initiate preschool screening programs. At the time of the BEP passage, five school systems had already established preschool screening programs. Three preschool screening approaches being used in the established programs were: the Gesell School Readiness Test (GSRT), the Developmental

Indicators for the Assessment of Learning-Revised (DIAL-R), and the Preschool Screening Battery (PSB) which consisted of the Developmental Test of Visual Motor Integration-Revised (DTVM-I), the Peabody Picture Vocabulary Test-Revised (PPVT-R), the Expressive One-Word Picture Vocabulary Test (EOWPVT) and the Preschool Screening Scale (PSS).

Numerous studies have been conducted on the psychometric qualities of these approaches (Ames & Ilg, 1964; Barnett, Faust & Sarmir, 1988; Duffey, Ritter & Fedner, 1976; Flynn & Flynn, 1978; Kaufman, 1971; Miller & Sprong, 1986; Suen, Mardell-Czudnowski & Goldenberg, 1989; Vacc, Vacc & Fogleman, 1987; Wood, Powell & Knight, 1984). However, despite the popularity of these screening procedures, there are few predictive validity studies in the professional literature.

These screening instruments have been used to identify young children with potential learning problems who might need special education placement or referral for a diagnostic evaluation (Suen, Mardell-Czudnowski & Goldenberg, 1989), and to identify preschool children who were considered to be too immature to attend kindergarten (Ames & Ilg, 1964). They have also been used to study the developmental aspects of early reading acquisition (Fletcher & Satz, 1980; Sears & Keogh, 1993; Stanovich et al., 1984).

Some researchers who focused on identifying the early determinants of reading acquisition concentrated on the effects of cognitive abilities (Stanovich et al., 1984), while others studied the neuropsychological correlates of reading achievement (Fletcher & Satz, 1980; Teeter, 1985). Tramontana, Hooper & Selzer (1988) reviewed 74 studies published from 1973 to 1986 on the prediction of reading achievement, and identified eight categories of predictor variables frequently used in this research. They were: readiness skills; general cognitive ability; specific cognitive abilities; language abilities; perceptual/ perceptual motor abilities; motor skills; behavioral-emotional functioning; and demographic factors.

Rationale for Study

This study was designed to help clarify which developmental abilities, measured by three preschool screening batteries, are most effective in predicting performance in specific areas of reading achievement at the third and fifth grades. Although numerous research studies have been conducted on the early prediction of reading achievement, the results have been mixed. Because of this ambiguity, clarification is needed in order to identify the determinants of early reading acquisition (Tramontana et al., 1988).

Adding to the confusion, the relationships between predictor variables and achievement measures have varied according to the type of outcome measures used and the grade level of the subjects at the time they were assessed (Tramontana et al., 1988). Developmental factors seem to be associated with these shifts in the predictive patterns of variables over time (Fletcher, Satz & Morris, 1984; Sears & Keogh, 1993; Stanovich et al., 1984).

Stanovich et al. (1984) found intelligence measures contributed significantly to the prediction of reading achievement at the third grade, but contributed negligibly to the prediction of reading performance at first grade and fifth grades. Jenson (1980) contended, however, that general intelligence was the single largest contributor to the prediction of reading achievement.

As young children with more obvious handicapping conditions are likely to be identified through early screening and diagnostic processes required by Public Law 99-457, preschool screening procedures geared toward the identification of youngsters with more subtle vulnerabilities to learning failure are needed. Increased environmental risks such as maternal drug and alcohol use, inadequate diet due to poverty, and chronic diseases which may affect the neuropsychological development of young children place them in the at-risk category (Roth, McCaul & Barnes, 1993).

Current knowledge regarding reading acquisition and the significant determinants of reading skill development may be enhanced by the research proposed in this study. Additional knowledge regarding the predictive validity of the screening tests will help school personnel to make decisions regarding their future use.

Statement of the Problem

Despite the increase in statewide developmental screening practices in North Carolina and federal legislation mandating early intervention, the need for preventive programs for young children at-risk for learning failure continues. Additionally, more information on the predictive validity the three test batteries selected for this study is needed to support their continued use.

Purpose of the Study

The purpose of this research is to compare the predictive validity of three preschool screening batteries utilized in North Carolina and to determine if the effects of early determinants of reading achievement measured at the time of school entrance significantly influence reading achievement. Review of the literature in Chapter II on reading acquisition and early determinants clearly confirms the need for more research in this area.

The present study is designed to answer these questions:

1. What is the predictive validity of the Gesell School Readiness Test (GSRT), the Developmental Indicators for the Assessment of Learning (DIAL-R), and the Preschool Screening Battery (PSB) to identify students at-risk for reading disability?

2. What are the early determinants identifiable through discriminant function analysis of these test batteries which contribute significantly to reading achievement across grades?

3. Does the effect of the determinants vary or change when they are associated with different outcome measurements such as word analysis, vocabulary, reading comprehension or total reading scores at third and fifth grades?

4. Are there significant differences in groups of children who were correctly identified as not at-risk by preschool screening data, and those who were classified as not at-risk by preschool screening data who later developed reading disabilities?

Limitations of the Study

The sample consists of students who were enrolled in five school systems in North Carolina during the 1988-1989 school year. Since all students in the sample resided in North Carolina at the time of preschool assessment, the results should be generalized to states with similar characteristics. The outcome measure for reading achievement at the fifth grade was also standardized on a sample of North Carolina students which limits generalization.

Attrition is another pitfall which influences longitudinal research (Fletcher et al., 1984). Due to the mobility of the student population, the original cohort of students who entered the target schools in 1988 most likely underwent change during the six-year interval from 1988 and 1994.

Selective sampling is another limitation to the study. Since informed consent is necessary for student record reviews, the make-up of the final sample will be affected by parental willingness to participate.

Definitions of Terms

The research outlined in this study is based on the following definitions:

(1) Developmental abilities are quantified measures of the pre-kindergarten child's level of functioning in cognitive, language, and motor areas as measured by the GSRT, DIAL-R, or PSB within the four month interval immediately preceding kindergarten entrance.

(2) Early predictors are developmental abilities measured by subtest scores from the GSRT, DIAL-R, and PSB which prove more effective than chance at predicting reading group placement at third and fifth grades.

(3) Early determinants of reading progress are those constructs or abilities that:

(a) emerge from interpretation of the functions extracted through discriminant function analysis of the three preschool screening batteries and other variables such as gender, ethnicity, socioeconomic status, and reading achievement.

(b) initiate a "causal chain of achievement" (Stanovich et al., 1984, p. 297) in which early developmental abilities act as precursors to acquisition of specific reading skills such as word analysis and vocabulary which in turn affect reading comprehension.

(4) Reading achievement is defined as total reading scores on the California Achievement Test, North Carolina End of Grade Developmental Reading scores, and reading grades at the third and fifth grade.

Summary

In summary, despite increased interest in preschool screening and prevention of school failure through early intervention programs, many children continue to fail in school. Improved procedures are needed for identifying the young child at-risk for reading failure.

This study is designed to examine the predictive validity of three preschool screening batteries to predict later reading achievement. The effects of developmental abilities on the early acquisition of reading skills in the primary grades and reading comprehension abilities at the end of elementary school were also studied.

CHAPTER II

Review of Literature

Interest in early identification and early intervention of potential learning problems may be inferred from the expansion of research and the growth of programs in these areas over the last three decades. Several factors contributed to this growth. The passage of Public Law 94-142 (The Education for All Handicapped Children Act) resulted in expanded efforts to identify school-aged children potentially at-risk for learning problems (Tramontana et al., 1988).

As a result, wide-scale screening and identification of children with learning disabilities occurred for the first time. Changes in the mental health field also contributed to this expansion. During this time, mental health professionals became more aware of the long-range consequences related to the unmet needs of learning problems on the mental health of school children (Satz & Fletcher, 1988). As a result of this heightened awareness, the President's Commission on Mental Health studied the effectiveness of early intervention and reported that early detection of learning disabilities was supported by "empirical evidence" (Satz & Fletcher, 1988, p. 824).

Advances in the field of infant and toddler behavior also generated new information regarding the perceptual and discriminatory abilities of infants and young children (Aslin, 1988; Gibson, 1969; Kochanek, Kabacoff & Lipsitt, 1990). This increased understanding of the "learning capacities of newborns and the role of perinatal risk in compromising growth and development" (Kochanek at al., 1990, p. 528) resulted in increased acknowledgement of the impact of life experiences on infants and young children.

Public Law 99-457 was passed in 1986 extending the right of a free appropriate education, guaranteed under Public Law 94-142, to children and their families from birth to five years of age. This legislation rekindled research interest in early identification and early intervention (Kochanek et al., 1990). Much of this research focused on early detection of the child with more obvious handicaps such as mental retardation or physical disabilities. Identification of the preschool child who was at-risk for subtle problems such as learning disabilities was more difficult. One of the requirements for identification of learning disabilities, as outlined in PL 94-142, is the presence of a discrepancy between achievement and ability. Clearly, preschool children who have not been exposed to formal school training in reading and writing will not exhibit a discrepancy between ability and educational achievement. As an alternative, Kirk (1987) argued for the inclusion of developmental learning disabilities, a category of learning disabilities in which the prerequisite skills necessary for learning academic subjects are affected.

Since the goal of early detection is to intervene in the early stages of reading acquisition, the capability to predict which children are at-risk for reading failure is paramount. The problem of prediction is attenuated by the fact that predictable precursors of reading performance have not been identified adequately at this time. Despite these difficulties, evidence to date does suggest that early identification and remedial intervention at the preschool level can be effective in reducing subsequent school failure (Satz & Fletcher, 1988; Tramontana et al., 1988). Meisels (1987) cited numerous researchers who have substantiated "the ameliorative effects of environmental intervention on childhood development" (p. 69).

The importance of reading performance in the acquisition of knowledge and the pursuance of adult roles is also well documented. Fletcher et al. (1984) discussed the plight of the adult disabled reader, and advocated for more research on early identification and prevention of reading disabilities. According to these authors, "investigators in recent years have unanimously expressed the need for a more valid early-warning screening program to begin in kindergarten before formal reading training begins" (Fletcher et al., 1984, p. 294).

Another group of researchers embraced the maturationist viewpoint in developing theory about early intervention. According to this theory, children were expected to learn to read when they were ready. Proponents of the maturational theory advocated for providing the developmentally-young child with additional time to mature before school entrance or before formal academic training begins (Ames & Ilg, 1964).

Other theorists argued that intervention should be commenced when the child is young and the nervous system is more responsive to change (Bloom, 1964; Caldwell, 1968; Satz & Fletcher, 1988). These writers theorized that the younger child is more malleable, a factor which would increase the probability of successful intervention.

In one early study, Strag (cited by Satz & Fletcher, 1988) found 82% of the children diagnosed as dyslexic by 2nd grade were successfully completing regular classroom work by the end of high school as opposed to 46% who were not identified as learning disabled until 3rd grade. Other studies have established the rationale for early identification (Bruck, 1985; Edgington, 1975; Muehl & Forell, 1975).

Finally, Roth et al. (1993) proposed that the predictive validity of fine motor and auditory skills at first grade may direct those interested in early intervention toward more extensive modification of instruction than is currently being done. They postulated that better planned interventions should lead to a decrease in school failure which, in turn, would result in diminished use of less desirable intervention practices such as retention.

Early Predictor Research

Before 1982, research involving the early prediction of reading achievement focused on univariate prediction comparing global ability with later achievement. Tramontana et al. (1988) reviewed 74 studies on preschool prediction of later school achievement and reported that approximately one-third of the investigations studied the validity of IQ scores as predictor variables for later achievement. Specific abilities such as memory, associative learning, and conceptual abilities were most often "embedded in batteries of tests" (Tramontana et al., 1988, p. 114). Those reviewers concluded it was unclear whether a specific cognitive ability, or a more general ability factor, had been responsible for the "obtained predictive relationship" (p. 114).

Other authors have called for more specification of predictor variables and reading subskills (Sears & Keogh, 1993; Stanovich, 1987), and in a number of studies cognitive abilities were divided into subdomains. When this design was used, verbal comprehension was often found to have high predictive validity (Stanovich, 1987; Stevenson & Newman, 1986).

Predictor Variables

Conceptual Variables

Stanovich et al. (1984) studied the relationship between general intelligence as measured by the PPVT-R and Ravens Progressive Matrices Test (RPM), decoding speed, phonological awareness, and listening comprehension scores of a group of 56 first-grade students. These authors found that all four measures were moderately related to end-of-year reading comprehension scores, but decoding speed accounted for the largest amount of unique variance. Comparisons across age levels indicated that the influence of general intelligence on reading achievement increases over time. Correlations between the RPM scores and the Metropolitan Reading scores at the first-, third-, and fifth- grades were: 0.33, 0.42, and 0.56. Correlations between the PPVT-R and the Metropolitan Reading scores at the same grades were: 0.34, 0.59, and 0.58.

Perceptual/Perceptual-Motor Variables

Over half of the studies reviewed by Tramontana et al. (1988) included some aspect of verbal and perceptual/ perceptual-motor abilities among the predictor variables. In several of these studies perceptual variables were found to be significant predictors of reading performance (Butler, 1985; Roth et al. 1993).

Perceptual/perceptual motor performance has been measured in different ways. Two measures commonly used were the Bender Visual Motor Gestalt Test and the Developmental Test of Visual Motor Integration. Other researchers used copying forms and human figure drawings to measure visual perceptual abilities and perceptual motor abilities (Duffey et al., 1976; Telegdy; 1985). Tramontana et al. (1988) concluded that the "greater weight of evidence" supported the belief that visual perceptual/motor measures do contribute significantly to the prediction of reading achievement (p. 115).

Preschool Screening Tests as Variables

Other research studies used subtest scores and total test scores from preschool screening batteries as predictor variables for later reading achievement (Barnett at al., 1988; Flynn & Flynn, 1978; Fletcher et al, 1984; Linguist, 1987; Rafoth, 1988; Roth et al., 1993; Stevenson & Newman, 1986; Wood et al., 1984; Vacc et al., 1987). Discriminant function analysis was used in some of these studies to reduce a large set of variables to a smaller number of more easily interpreted factors. In a few cases, experimental tests were used to measure specific variables.

Roth et al. (1993) administered the Early Prevention of School Failure Battery (EPSF) which includes five screening tests, the PPVT-R, DTVMI, Preschool Language Scale, Draw-A-Person and Revised Motor Activity Scale, to 161 kindergarten students. Discriminant function analysis was then conducted to determine how effective those predictors were in identifying students who were at-risk for retention, or who had been referred for special education evaluation and placement.

These researchers also conducted regression analysis using the screening subtests as predictors of reading scores. The auditory variable (a combination of the Preschool Language Scale and the PPVT-R) was a significant predictor of fall reading scores when age, gender, and socioeconomic status were controlled. Auditory and fine motor predictors were significant in predicting spring reading.

Socioeconomic Variables

Socioeconomic status (SES) and gender were included as they have been discussed extensively in professional literature. Tramontana et al. (1988) reported most studies reviewed included information regarding demographic factors, but less than one-half included these factors in their prediction analysis. When gender was included, girls appeared to out-perform boys. Mixed results were reported when SES was included as a predictor variable. Studies by Evans (1977), Gold & Berk (1979) and Kohn & Rosman (1974) (cited by Tramontana et al., 1988) found lower SES was associated with poorer outcomes. In contrast, Abrahamson & Bell (1979) found predictive accuracy to be lower with high SES children. Finally, Tramontana et al. (1988) reported results from two studies, Mendels (1973) and Kohn and Rosman (1974) in which the inclusion of demographic variables provided a "small but significant increase in predictive accuracy" (p. 118) when added to cognitive measures.

Sears & Keogh (1993) found that early identification appears to be more predictive for girls and middle class children. When these researchers compared measures obtained at kindergarten with reading performance at first and third grades, they found gender differences in the patterns of correlations between kindergarten abilities and later achievement. Correlations for girls ranged from r = 0.52 to

0.33 in comparison with correlations ranging from r = 0.46 to 0.14 for boys. In general, girls scored higher than boys on all measures except visual skills. Correlations comparing preschool visual skills with word study skills and reading comprehension were 0.32 and 0.31 respectively for boys. Comparable correlations for girls were 0.19 and 0.05. Sears & Keogh (1993) concluded that the "routes to reading competency may be somewhat different for boys than for girls" (p. 85).

Research by Vacc et al. (1987) found that SES and race contributed significantly to reading performance as did Stevenson and Newman (1986). These researchers also found gender differences in a longitudinal study in which prekindergarten measures were used to predict achievement up through the tenth grade.

On a broader scope, Yesseldyke and O'Sullivan (1987) studied the contribution of demographic and educational variables to the prediction of screening referral rates in a statewide preschool screening program. Principal components analysis was conducted and five global factors were extracted. Factor I was labeled SES and had high loadings on measures such as median income and median rent. Factor II was labeled school district size, and Factor III was labeled minority/federal review and had high loadings on measures such as number of minority students and results of federal reviews. The other two factors dealt with funding and spending patterns of the school districts.

Stepwise multiple regression analysis was computed to predict screening referral rates from the demographic/ educational variables. The results indicated that preschool screening outcomes such as referral rates were not related to broad socioeconomic and educational factors. The authors acknowledged, however, that the results of their study were relevant to larger social systems and might not be relevant to individual children and their families.

Longitudinal Research

Fletcher et al. (1984) conducted a longitudinal study aimed at identifying developmental precursors to academic achievement with emphasis on reading performance. Their project included standardization and validation of a kindergarten screening battery which consisted of the PPVT, the DTVMI, alphabet recitation, finger localization, a somatosensory test, a shortened version of the Wepman Auditory Discrimination Test, verbal fluency tasks, the Similarities Subtest of the Wechsler Intelligence Scale for Preschool and Primary Scale (WPPSI), and tasks which measured neuropsychological functioning. Achievement criteria were comprised of teacher and test-based assessments. Teacher-based criteria were obtained from end-of-year classroom reading levels, and test-based criteria were obtained from the Iota Word Test. Discriminant function analysis was computed after all test tasks were entered, and the following factors were extracted: Factor I - Sensorimotor-perceptual, Factor II - Verbal-Conceptual, and Factor III - Verbal-Cultural. Four predictor variables accounted for most of the variance in reading achievement. The final screening battery was comprised of the following predictors: the PPVT, the alphabet recitation subtest, the recognition-discrimination subtest, and the DTVMI. This battery was then used for cross-validation with other samples, and was found to identify the majority of children who became disabled readers by fifth grade. The overall hit rate was 0.69 with minimal classification errors with average to superior readers (Fletcher et al., 1984).

Sears and Keogh (1993) examined the relationship between the Slingerland Pre-Reading Screening and performance on word recognition and reading comprehension. In this study, 104 children were administered the Slingerland Procedures at kindergarten and tested again with the word recognition and comprehension subtests of the Stanford Achievement Test at first, third, and fifth grades. Five subtests of interest were identified based on reading theory (Sears & Keogh, 1993). Test 2 consisted of visual processing tasks which required sensitivity to letter order and reversals, Tests 5 and 8 measured listening comprehension, and Test 12 measured phonological awareness. The last variable, Test 6, measured letter-name knowledge.

Sears and Keogh (1993) found overall Slingerland test scores to be significantly related to later reading performance. Specific subtests differed in strength of association and varied across grade levels. The predictive antecedents which emerged for word recognition were measures of phonological awareness, letter-name knowledge, and visual matching, whereas, listening was related to comprehension. One may conclude from these results that the antecedent skills needed in the early stages of reading acquisition are different from those needed at later stages of reading comprehension. Sears and Keogh (1993) advocated for future research on the predictive relationships between preschool abilities, decoding, and reading comprehension.

Fletcher and Satz (1980) studied the developmental changes in the neuropsychological correlates of reading achievement. Canonical Variate Analysis was conducted to "determine the contribution of different performance dimensions to reading group discrimination across age" (Fletcher & Satz, 1980, p. 30). The results of this study indicated that measures of sensorimotor-perceptual skills (finger localization and visual discrimination) contributed more to the predictive validity between preschool abilities and reading achievement at the second and fifth grades than verbal and conceptual skills.

Two other predictive categories identified by Tramontana et al. (1988), motor skills and behavioral-emotional functioning, will not be discussed extensively. Tramontana et al. (1988) reported motor skills (perceptual-motor abilities excluded) had received little attention as predictor variables, and concluded that fine and gross motor skills did not appear to be one of the better predictors of later achievement. They did not, however, define fine motor skills nor describe how they were measured.

Preschool Screening Batteries

The Preschool Screening Batteries to be studied in this research will be discussed in this section. Each battery will be described, and validity and reliability data will be presented. Research in which the screening tests were used as predictor variables of later achievement will be discussed.

Developmental Indicators of Learning Assessment-Revised (DIAL-R)

The DIAL-R is a popular preschool screening test used to classify children into three categories: potential problem, "ok", and potentially gifted. Cut-off scores for the identification of the potential problems and the potentially gifted groups is 1.5 standard deviations from the mean. Children who fall into the potential problem group are recommended for further evaluation. Norms tables provide the following scores for three-month age intervals: means, standard deviations, standard errors, and internal consistency reliabilities (Poteat, 1989).

The normative sample consisted of 2,447 subjects with males and females being equally represented. Some difficulties were noted with sample stratification suggesting that results from culturally and socio-economically different samples be interpreted cautiously (Poteat, 1989).

Content validity for the battery, which consists of Motor, Concepts, and Language subtests, was established from interviews with teachers and early childhood specialists. Factor analysis revealed that the Motor and Concepts subtests correlated highly with each other and resulted in one un-named factor. The second was considered to be a language factor. Correlations among the three subtests ranged from 0.77 to 0.85. Based on this analysis, use of the total test score was recommended (Poteat, 1989).

The DIAL-R was also compared with the Stanford-Binet to obtain criterion validity. A correlation of 0.40 was obtained between the DIAL-R and Stanford-Binet scores. The two scales agreed 82% of the time on classification decisions based on the DIAL-R screening categories.

Content validity was established in the standardization of an earlier version, the DIAL, through interviews with teachers and early childhood specialists (Sattler, 1992). Criterion-related validity was measured by correlations between the Stanford-Binet IQ scores and the DIAL-R total and subscale scores of 125 children. They were: 0.40 Total Test, 0.33 Language, 0.50 Concepts, and 0.28 Motor (Poteat, 1989). Information on predictive validity was not found in a literature search according to Poteat (1989).

Test-retest correlation coefficients were presented for 65 children who were retested after a two-week period. They were: Motor ($r = 0.76$), Language ($r = 0.77$), Concepts ($r = 0.90$), and Total Test, ($r = 0.87$) (Poteat, 1989). Cronbach's alpha was 0.96 for the total scale. Poteat (1989) reported the DIAL-R seems to have adequate internal consistency with the exception of the Language Subtest which has a "pattern of declining internal consistency associated with increasing age" (p. 247).

Lichtenstein (1981) conducted a comparative validity study between the DIAL and the Denver Developmental Screening Test and three criterion measures: The Stanford-Binet Intelligence Scale, the PPVT, and the Woodcock-Johnson Psycho-Educational Battery. Correlational analyses were computed between the two screening batteries and outcome measures. A correlation of 0.82 was obtained between the total test scores of the two batteries, but the tests were in complete disagreement in classification outcomes thus lacking concurrent validity. The correlation between the DIAL and the PPVT was 0.52 when age was controlled.

Miller & Sprong (1986) conducted a qualitative comparison of four preschool screening instruments which included the DIAL-R. These authors reported that the DIAL-R met most of the ten psychometric criteria used in the study.

Several studies in which the DIAL or the DIAL-R were used in preschool screening research were published during the Eighties. Vacc et al. (1987) conducted a predictive study comparing the original DIAL with first grade performance on the California Achievement Test (CAT). A group of 254 entering kindergarten students were administered the DIAL in May and August, and the California Achievement Test in the Spring of the first grade. Canonical correlation analysis was computed to determine the relationship between DIAL scores and achievement measures.

Canonical correlation analysis was conducted to assess the relationship between the four DIAL subtests and the four CAT subtests. The Concepts and Fine motor subtests were the most significant predictors of reading performance at the first grade level with correlations of 0.35 and 0.38

respectively. Race and SES were also included in the Best model.

Barnett, Faust, and Sarmir (1988) conducted a validity study of two preschool screening instruments, the DIAL-R and the Learning Accomplishment Profile-Diagnostic (LAP-D). The intent of this study was to examine the convergent validity of two preschool screening instruments using the multitrait-multimethod model. Like-named scales on the DIAL-R and the LAP-D were compared through regression analysis. The analysis resulted in a correlation of 0.86. The Cognitive subtest on the LAP-D also correlated at 0.86 with the Motor Scale on the LAP-D. Because of these discrepancies, profile interpretation of these batteries was not recommended (Barnett, et al., 1988).

Internal consistency reliabilities measures reported ranged from 0.41 to 0.88 across ages levels and subtests. Suen at al., (1989) reanalyzed the norming data of the DIAL-R using the squared-error loss generalizability approach to assess classification reliability of subtest and total test scores. This analysis resulted in Cronbach Alphas for total scores above 0.95. Using the cut-off of 0.80 as an acceptable measure of reliability, as suggested by Salvia and Yesseldyke (1985), 83 percent of the subtest reliabilities were below 80. On the other hand, these authors contended that the decision reliabilities of the

DIAL-R are quite high when one considers the Cronbach Alphas (0.95) associated with total scores (Suen et al., 1989).

In a review of 29 preschool screening tests conducted in 1988, Witt, Elliott, Gresham and Kramer (cited by Suen et al., 1989) concluded that the original DIAL was among three of the most frequently used screening batteries, and was one of the most comprehensive and psychometrically sound tests available at the time. These authors also concluded that the DIAL-R, the revised version of the original DIAL, was vastly improved.

Gesell School Readiness Test

The Gesell School Readiness Test (GSRT) was constructed from tasks developed by the Gesell Institute of Child Development. The behavioral tasks constructed by Gesell were based on Gesell's maturationist viewpoint (Meisels, 1987). From this viewpoint, behavior is considered to be a "function of structure, changing in patterned predictable stages" (Meisels, 1987, p. 69).

In a critique of the GSRT published in the *Ninth Mental Measurements Yearbook*, Bradley, (1985) cited the following weaknesses: outdated norms, restrictiveness of the standardization sample, lack of reliability and validity, and unclear scoring procedures. Strengths noted were: well-designed, familiar, test materials, and "very valuable descriptive material" (Bradley, 1985, p. 611).

Validity was established through comparisons of Gesell scores with teacher ratings of student performance at the end of year. A high level of agreement was reported between GSRT scores and sixth-grade performance (Bradley, 1985). Bradley cautioned, however, that extreme scores may have influenced this association. Little information on reliability is provided in the manual.

Louise Bates Ames and Frances Ilg, two co-workers at the Gesell Institute, adopted the concept of developmental age and theorized that school readiness should be based on behavioral maturity rather than chronological age (Ilg & Ames, 1965). According to their philosophy, children should be placed in school according to their behavioral functioning rather than their chronological ages (Ames & Bates, 1964). Although the Gesell Institute claims the test is based on fifty years of experience, Kaufman (1971) stated that investigations of the "relationships between performance on the GSRT with other variables has been scanty" (p. 1343).

Samuel Meisels (1987), a critic of Gesellian developmental theory, conceded that the GSRT has increased in popularity over recent years. According to Meisels, the Gesell Institute asserted that "thousands of public, private, and parochial schools nationwide" (p. 69) have adopted the Gesell developmental tests. Despite this popularity, Meisels criticized the paucity of validity studies on the GSRT reporting that the Gesell Institute had no validity studies available as of March 28, 1984. Meisels stated that representatives of the Gesell Institute explained to him, through personal communication, that the Gesell Test "documents normative responses by age" (Meisels, 1987, p. 70), and that validity had been established through years of experience and use.

Despite this criticism, there is some evidence of predictive validity. Ames and Ilg (1964) compared chronological age at school entrance, readiness for kindergarten as measured by the Gesell Behavior Test, two projective tests (the Rorschach and the Mosaic Tests), and a vision examination with actual sixth grade school performance. Readiness for first-grade work was determined by independent judges. At sixth grade, the children were divided into four groups on the basis of the quality of their sixth-grade work (group I membership denoted highest performance). Of the children rated as not ready for kindergarten in 1957, all had failed a grade or had been placed in group 3 indicating poor sixth grade performance.

Wood et al. (1984) studied the predictive validity of the GSRT by comparing preschool scores with kindergarten performance and referral for special needs evaluation. Any child who experienced adjustment difficulties during kindergarten was referred to the special needs committee. Subsequent discriminant analysis was performed using the Gesell developmental age and chronological age as predictors. The two groups were designated as special needs versus nonspecial needs. Differences between the two groups on developmental age were significant while differences between chronological age were not. Developmental age accounted for "22% of the variance of the criterion measure" (Wood et al., 1984, p. 9). Chronological age did not add significantly to prediction. These authors concluded that a developmental age of 55.6 months or 4.6 years is a critical cut-off for identifying children ready for kindergarten. They contended that their research supported the validity of the GSRT.

In order to answer some questions regarding the psychometric properties of the GSRT, Kaufman (1971) conducted a study in which the GSRT and a battery of tasks built upon Piaget's tasks were compared with the Lorge Thorndike Intelligence Test. In this study, a sample of 103 beginning kindergarten students were administered the GSRT in September. During the school year, the children were also given the Lorge-Thorndike Intelligence Tests and a battery of Piaget's Tasks. Kaufman (1971) reported reliability coefficients of 0.80 for females and 0.84 for males for the GSRT. The mean score for females was 50.2 and 40.5 for males. Kaufman (1971) suggested that reliability coefficient alpha's of 0.80, and the standard error which averaged 5 to 6 points made the GSRT "reasonably reliable" (p. 1352).

Factor analysis of the GSRT was performed to extract underlying constructs. Three factors emerged after Verimax rotation. Factor I was interpreted as a paper and pencil coordination construct, and all tasks loading on this factor involved motor coordination and pencil manipulation.

Factor II was labeled as awareness of part-whole relationships and included two incomplete man tasks. Finally, Factor III was labeled as academic achievement, and involved the ability to write (an experience-dependent skill). Kaufman (1971) concluded that the GSRT might be interpreted as "two parts behavioral maturity (Factor I), and one part abstract intelligence (Factor II), and one part experience (Factor III)" (p. 358). Kaufman offered an expanded interpretation of the Gesell in which behavioral maturity is only one of three factors measured by the test. This differed from the narrow interpretation espoused by Ilg & Ames (1964) who theorized the Gesell School Readiness Test measured a single construct, behavioral maturity (Kaufman, 1971).

In maturationist theory, behavior is viewed as a "function of structure, changing in a patterned, predictable way" (Meisels, 1987, p. 69), and precludes educational and environmental interventions as time is considered the crucial variable. The Gesell Institute advocates for using developmental age to determine when a child is ready for placement in kindergarten or first grade. Various studies have examined the effects of developmental placement and early retention (Ames & Ilg, 1964; May & Welch, 1984; Wood et al., 1984). Studies conducted by Ames and Ilg (1964) and Wood et al., (1984) reported positive results for children placed on developmental age. May and Welch (1984) did not find "demonstrable positive effects of the practice of 'buying a year' on children's later academic achievement" (p. 385).

Preschool Screening Battery

The Preschool Screening Battery consists of four separate instruments which will be discussed separately. Three of the tests, the Peabody Picture Vocabulary Test-Revised (PPVT-R), the Developmental Test of Visual Motor Integration-Revised (DTVMI), and the Expressive One Word Picture Vocabulary Test (EOWPVT) are nationally normed tests. The Preschool Screening Scale was developed and normed locally.

Peabody Picture Vocabulary Test. Scores from the PPVT-R are often used as measures of receptive language or verbal intelligence. Split-half reliability coefficients reported in the manual range from 0.61 to 88. Alternate form reliabilities range from 0.71 to 0.91 for the standardization sample (McCallum, 1985). Validity studies which have correlated the PPVT-R with intelligence measures using the WISC-R have yielded correlations from 0.16 to 0.86 (Sattler, 1992). Test-retest reliability was computed on a subsample of 962 children who were tested within a time-interval ranging from 9 to 31 days. Coefficients for standard scores ranged from 0.54 to 0.90 with a median coefficient value of 0.77 (McCallum, 1985).

Reliability of performance was judged adequate over short-time intervals. Stability coefficients for the age ranges, 2 years-6 months, to 2 years 11 months, 5 years to 8 years-11 months, and 18 years to 18 years-11 months were relatively low. Validity studies were reported to support content, construct, and concurrent validity. Content validity was provided by item selection procedures such as vocabulary searches and cross-referencing with age and grade-level vocabulary lists. Construct validity was substantiated by an increase in number of correct items reported as age increased. In a critique of the PPVT-R published the *Ninth Mental Measurements Yearbook*, McCallum,

(1985) described the psychometric characteristics of the PPVT-R to be "adequate to excellent" (p. 1127).

The PPVT-R has been administered separately or as part of a preschool screening battery in numerous studies (Butler, Marsh & Sheppard, 1985; Colarusso, Gill, Plenkenhorn & Brooks, 1980; Fletcher et al., 1984; Flynn & Flynn, 1978; Lewis, 1980; Roth et al., 1993; Stevenson & Newman, 1986). Tramontana et al. (1988) reported mixed results regarding the performance of the PPVT-R as a predictor variable. Their conclusions supported the use of the PPVT-R as a predictor of reading in the later grades but not as a predictor of overall achievement especially in samples with high proportions of minority students.

The PPVT-R has been interpreted to measure receptive language and intelligence levels in the following studies. Flynn & Flynn (1978) administered the PPVT as part of a preschool screening battery to 81 children. Stepwise regression analysis was conducted, and these authors reported a 0.004 correlation between PPVT scores and the reading score on the California Achievement Test at the second grade level. These findings confirm later research in which language scores were more predictive of reading achievement in comprehension at upper elementary level than in early acquisition of reading (Stanovich et al., 1984).

Other studies found the PPVT-R loaded most heavily on a factor labeled verbal reasoning (Fletcher & Satz, 1980; Fletcher et al., 1984.) Fletcher and Satz (1980) performed canonical variate analysis and found that PPVT-R scores and the Similarities subtest correlated highest with the canonical variate at R = 0.69 and 0.74 consecutively at the fifth grade. At the same time, concurrent validity of areas such as sensiormotor-perceptual measures declined.

In another study, Stanovich et al. (1984) studied the relationship between intelligence, cognitive skills and early reading progress. Questioning the hypothesis that reading comprehension is highly related to a matter of general intelligence, these researchers examined the relationships between "four determinants of reading comprehension: phonological awareness, decoding speed, listening comprehension, general intelligence, and reading comprehension" (Stanovich et al., 1984). Raven's Progressive Matrices (RPM) and the PPVT-R were administered to all subjects. The RPM was considered to be a culture-reduced measure while the PPVT-R was considered to be a measure of general intelligence. The PPVT-R correlated (R = 0.33) significantly with listening comprehension at the first grade. Loadings on the RPM and PPVT-R did not converge as expected. At the third-grade level, the PPVT-R had the highest correlation with the Metropolitan scores.

When RPM and PPVT-R were combined at first-, third- and fifth-grades, they explained 18.6%, 36.7%, and 42.5% of the variance in reading comprehension respectively. These authors concluded that the relationship between intelligence and listening comprehension measures increased by third grade. Also, correlations between PPVT-R, RPM, and listening comprehension increased from first to third and from third to fifth.

Similarly, Stevenson & Newman (1986) obtained the following zero order correlations between the PPVT-R and fifth grade decoding scores on the Wide Range Achievement Test and the Gray Reading Comprehension Test at fifth grade (0.19 and 0.43) respectively. These findings substantiated the hypothesis that language abilities, especially receptive language, are more important to acquisition of reading comprehension abilities than decoding skills and early acquisition skills.

Roth et al. (1993) studied the Early Prevention of School Failure (EPSF) procedure in which the PPVT-R is combined with the Preschool Language Scale, a criterion-referenced test which assesses "integrative and conceptual areas", into an auditory modality variable (p. 350). Roth et al. (1993) studied screening results of the EPSF battery and found the auditory modality to be a significant predictor of fall reading achievement when age, gender, and SES were controlled.

Developmental Test of Visual Motor Integration. The DTVMI is a frequently used screening instrument which measures pencil manipulation and fine motor development. Reliability measures reported in the original manual were based on 1964 norms, and test-retest reliabilities ranged from 0.63 to 0.92 with a median correlation of 0.81 (Beery, 1989). Time intervals ranged from 2 weeks to 7 months. Internal consistency reliabilities ranged from 0.66 to 0.93 with a median of 0.79. Interrater reliabilities ranged from 0.58 to 0.99 with a median of 0.93 (Beery, 1989). Validity measures, based on the 1964 norms, reported correlations of 0.89 between DTVMI scores and chronological age. The 1964 sample was comprised of 1,030 children from urban, suburban, and rural Illinois. In 1967, norms were developed for children of various minority groups such as Chinese and Greek nationalities. Cross-validation was conducted in 1981 using the 1967 norms with a sample of 2,060 subjects from a diverse population in California. The 1981 norms were found to be identical to the 1964 sample. The 1981 norms were subsequently added to the 1964 norms and published in the 1982 edition of the manual. Additional cross-validations were conducted in 1988 with 2,734 children from eastern, northern, and southern states. According to Beery (1989), the 1989 results were not significantly different from other norming groups.

Tramontana et al. (1988) cited seven studies from 1973 to 1986 which used the DTVMI as a variable for prediction of later achievement (Colarusso et al., 1980; Duffy et al., 1976; Fletcher & Satz, 1980; Flynn & Flynn, 1978; Klein, 1978; Lindgren, 1978; Reynolds, Wright & Wilkinson 1980). Flynn and Flynn (1978) included the DTVMI in stepwise regression analysis to predict reading scores on the Metropolitan Readiness Test at the end of kindergarten and the California Achievement Test at the end of second grade. The decreasing importance of visual motor abilities in the acquisition of reading ability from kindergarten to later grades was indicated by correlations of 0.317 on the Metropolitan Readiness Test and 0.019 with the California Achievement Test Reading subtest score at the second grade.

Duffey et al. (1976) studied the relationship between the preschool scores on the DTVMI, the Goodenough Draw-A-Person Test and second grade Stanford Achievement Test scores. Significant correlations were found between the DTVMI and the Stanford Achievement Test Scores: Vocabulary (29), Reading Comprehension (27) and Total Reading (28). DTVMI scores and total Stanford Achievement Test Scores were not significantly related.

Reynolds, Wright, and Wilkinson (1989) studied the incremental validity of the Test for Auditory Comprehension of Language and the DTVMI in predicting second-grade achievement, and both showed moderate correlations with the achievement measures (Univariate r's for DTVMI and Reading and Language Arts were 0.45 and 0.47 respectively). Although increases resulting from the addition of one measure with the other in the regression analysis were statistically significant, the authors concluded the increases which ranged from 3 to 8% were not of "practical utility" (Reynolds et al., p. 506).

Fletcher and Satz (1980) and Fletcher et al. (1984) included the DTVMI in a battery of preschool screening tasks which were entered into discriminant function analysis to determine membership in reading groups at grades 2 and 5. The DTVMI loaded most heavily on Factor I (which also had high loadings on finger localization and visual discrimination tasks). This factor was interpreted to be represent a sensorimotor-perceptual construct. Factor I accounted for 57% of the variance of predictive validity at grade 2, and 63% of the variance in reading performance at grade 5.

In the Roth et al. (1993) study, the DTVMI was combined with the Draw-A-Person Test to form a fine motor modality. These researchers found fine motor ability to be a significant predictor of spring reading scores. In conclusion, Roth et al. (1993) found the fine motor and auditory variables to be the best and most consistent predictors of future status related to retention and referral to special education (p. 356).

<u>Expressive One-Word Picture Vocabulary Test.</u> The EOWPVT measures verbal expression of language by assessing the child's ability to make associations between pictures and words. The test consists of 110 drawings which the examinee is asked to describe in one word. This test was developed by Morrison F. Gardner who stated the purpose of the EOWPVT is to estimate the expressive vocabulary of the child, and quality and quantity of vocabulary acquired from home and any formal education (Cummings, 1985).

The test was normed on 1,607 children ages 2-0 to 11-11 who lived in the San Francisco Bay area. The PPVT and the Columbia Mental Maturity Scale (CMMS) were used as anchor measures to assure adequate representation of nine IQ categories. Problems with the anchoring procedure were discussed by Cummings. The PPVT was used for ages 2-0 through 3-5 and the CMMS was used for the remaining ages. The CMMS was used for ages 4-0 through 6-0 which is the age interval of interest to this study. The CMMS norms were obtained in 1959 and the task requirements for the CMMS were discrimination and classification versus hearing vocabulary measured by the EOWPVT and the PPVT.

Split-half reliability coefficients ranged from 0.87 to 0.96. Cummings (1985) concluded that, at face value, the EOWPVT could be considered as adequate for providing a consistent estimate of expressive vocabulary. Specific information on how the split-half reliabilities were calculated was missing, however.

Two types of validation were reported. Content validity was established through questionnaires completed by parents to determine frequency of use of test items. Cummings (1985) concluded that the selection and representativeness of items were adequate, and the concepts represented appeared to be those likely to be encountered across diverse "socio-cultural backgrounds" (p. 565).

Criterion-related validity was computed through correlations with the PPVT (0.67 to 0.78), CMMS (0.29 to 0.59), the WPPSI (0.73) and the EOWPVT. The correlation between the EOWPVT and the WPPSI should be considered inflated because the author used mental ages rather than standard scores for this computation. Another caution regarding use of the EOWPVT scores is related to the narrow verbal functioning assessed by the test. As success on the test requires a one-word response, more elaborate verbal fluency is not assessed. Cummings (1985) concluded the

EOWPVT is best used as a screening instrument or as a supplemental test to be used as part of a language assessment battery.

Scoring the EOWPVT is quick and objective according to Spivack (1985). Raw scores are converted into Mental Age, Deviation Intelligence Quotients, Percentiles, and Stanine rankings. Mental ages below 2.0 and above 11.11 were extrapolated which facilitates use with exceptional children who might perform below or above the basal or ceiling scores computed for the standardization sample. Spivack (1985) concluded the EOWPVT could be appropriately used to assess school readiness when complemented by such measures as the PPVT, the McCarthy Screening Test, or the Bender-Gestalt Test.

Preschool Screening Scale. The PSS was developed in one of the target school systems. School psychologists, speech therapists, special education teachers, kindergarten teachers, occupational therapists, and physical therapists were involved in the development of the PSS. Content validity was accomplished through interviews with early childhood teachers, speech therapists, school psychologists, and occupational therapists. Reliability estimates were computed from the 1988 administration of the Preschool Screening Battery. Split-half reliability coefficients reported ranged from 0.7497 to 0.7565. A Spearman Brown

Reliability Coefficient of 0.7565 was also reported. Instructions are standardized, and all evaluators are trained in administration of the test. Scoring guides are included for individual items, and frequent responses are coded as plus or minus scores. Local percentile ranks for total Preschool Screening Scale scores are provided. A percentile score of 25% was established as a cut-off score for at-risk status. The PSS was developed for use by classroom teachers and other educational personnel.

Socioeconomic Status Index

Duncan's Socioeconomic Index (SEI) was devised to provide a prestige scale to all occupations listed in the 1950 Census. It has undergone revisions in 1960, 1970, and 1980. Scores for occupations ranged from 3 for motorman to 96 for dentists on the 1970 revision (Stricker, 1988). The original SEI scores were derived from the prestige ratings of occupations in the 1947 North and Hatt National Opinion Research Center (NORC) data (Duncan, 1961). The ratings of occupations were correlated with education and income data from the 1950 Census data. Ratings were estimated for all 496 occupations included in the Census (Stricker, 1988). In subsequent years, ratings have been assigned to additional occupations represented in the 1960 Census (Blau & Duncan, 1967), the 1970 Census (Hauser & Featherman, 1977), and the

1980 Census (Stevens & Cho, 1985). According to Duncan (1961) SEI scores can be estimated for any occupation for which education and income data are available.

Stricker (1988) used information from an open-ended questionnaire administered to 3,266 white and 202 black navy recruits to study the validity of the Duncan SEI and the Seigel Prestige Scale. Subjects were asked to categorize the occupation of the person who supported them before they left home. Occupational groups were assigned an SEI score and a Seigel Prestige Scale score. Other variables examined were main support's education in years, subject's education in years, a verbal ability measure, and an arithmetic reasoning score. Correlations were computed among the variables. The SEI scores and the Seigel Prestige scores for main support's occupation correlated highly with each other (0.92 for whites and 0.97 for blacks). The SEI and Seigel scores correlated with main support's years of education at 0.52 for whites and 0.48 for blacks (Stricker, 1988).

Caston (1989) compared scores on four dimensions of occupational inequality with SEI scores to determine if SEI alone is adequate to represent the "stratifying structures of American occupations" (p. 335). Caston used data collected by Blau and Duncan (1967) on 17,447 white males between the ages of 20 and 64. Variables included were:

Seigel Prestige Scale score, self-direction scores, 1960 medium income estimates (1960 U. S. Census), measures of educational and vocational preparation, and Duncan's SEI's. Results of the study indicated that the last three variables exerted more influence in the stratification of American occupations then prestige. Caston concluded Duncan's SEI provides a "good proxy for multi-dimensions of occupational inequality" (p. 344).

Achievement Tests

California Achievement Test Battery

The California Achievement Test Battery is a norm-referenced nationally validated set of achievement tests which measure three aspects of reading performance at the third grade level: word analysis, vocabulary, and total reading. Kuder-Richardson internal consistency reliabilities were reported in the high 0.80s to 0.90s for children tested above first grade (Airasian, 1989). According to Airasian (1989), reliabilities for younger children were lower (range 0.60s - 0.70s). Airasian (1989) also reported stability reliabilities for the elementary grades ranged from 0.80 to 0.95 and equivalent forms reliabilities for reading were "around 0.85" (p. 128). Vacc et al. (1987) found the concepts and fine motor subtests of the DIAL to be the most significant predictors of reading performance on the California Achievement Test at the first grade level with correlations of 0.35 and 0.38 respectively.

The North Carolina End-of-Grade Reading Test

The North Carolina End-of-Grade Reading Test is a test developed by the state of North Carolina to measure student progress on a continuous scale. In the manual, reliability measures were reported in standard error scores which represented a 2-3 point range for most test-takers with larger intervals (4-6 point range) occurring with extreme scores (Pommerich, Belleaud, Williams & Thisson, 1993). Test validity was measured by the relationship between teacher classified achievement levels for students, and the average scores of the students within each of those levels.

In 1992, teachers classified students into one of the four achievement levels. Students who were rated as high achievers by the teachers also obtained high scores on the end of course test, and conversely, students who had lower classifications scored lower on the end-of-course test. The teacher classifications were used to select cutpoints for classification of student performance on the 1993 reading tests, and each student was subsequently assigned an achievement level for the 1993 tests.

Summary

The purpose of this chapter has been to summarize early predictor research, and to provide psychometric data on the screening tests, reading performance measures, and the socioeconomic index to be used in this study. Factors related to increased interest in early prediction of reading disability, and issues related to the importance of early intervention were reviewed.

CHAPTER III

Methodology

The methodological procedures used in this study are summarized in this chapter. The research design is presented along with a restatement of the research questions. The population for the study is described along with the process used for the selection of subjects. The instruments used in measuring predictor variables and outcome measurements are identified and described. Finally, data collection and methods of data analysis to answer the research questions are discussed.

Research Design

The predictive validity of three preschool screening batteries was computed using a sample of sixth-grade students in five local educational agencies in North Carolina. The screening batteries were: the Gesell School Readiness Test (GSRT), the Developmental Indicators of the Assessment of Learning-Revised (DIAL-R), and the Preschool Screening Battery (PSB). Predictor variables from the three screening batteries were correlated with measures of reading performance at the third and fifth grades.

This research design was devised to address the following research questions.

Research Questions

1. What is the predictive validity of the Gesell School Readiness Test (GSRT), the Developmental Indicators for the Assessment of Learning (DIAL-R), and the Preschool Screening Battery (PSB) in identifying kindergarten students at-risk for reading disability?

2. What are the early determinants identifiable through discriminant function analysis of these test batteries which contribute significantly to reading achievement across grade levels?

3. Does the effect of the determinants vary or change when they are associated with different outcome measurements such as word analysis, vocabulary, reading comprehension, or total reading scores at the third and fifth grades.

4. Are there significant differences in groups of children who were correctly classified as not at-risk by preschool screening data, and those who were classified as not at-risk who later developed reading disabilities?

Participants

Participants consisted of sixth-grade students who participated in preschool screening as kindergarten students in five public school systems in North Carolina during May and August, 1988. School system selection was based on the type of screening battery used in each unit during the 1988-1989 school year, and the availability of screening data.

The screening batteries were selected to represent the most frequently used screening approaches in North Carolina during 1988. The North Carolina State Department of Public Instruction provided a list of 23 preschool screening incentive grants which were funded through Basic Education Program appropriations during the 1988-1989 school year. Procedures most frequently used by the 23 school districts were: the Developmental Indicators of Assessment in Learning-Revised (DIAL-R), the Gesell School Readiness Test (GSRT), and the Preschool Screening Battery (PSB).

All school systems on the list which had utilized one of the three screening procedures were contacted. One system declined to participate, and in two other systems, the preschool screening data had been purged from the records. In other cases, individual schools within a system had employed one of the screening procedures, but the total school system had not. In these cases data were not available for a representative sample of the system population.

Table 1 lists characteristics of the five school systems included in the study.

Table 1

School System Characteristics

Characteristic	DIAL-R School System 1	DIAL-R School System 2	PSB School System	GSRT School System 1	GSRT School System 2
Local Expenditures Per Student	$3695	$1091	$679	$1060	$750
Total Expenditures Per Student	$541	$4483	$3739	$4211	$3806
Average Local Supplement	$703	$1046	$ 0	$768	$596
Ethnicity					
American Indian	1%	1%	1%	1%	1%
Asian	1%	2%	1%	1%	1%
Hispanic	1%	1%	1%	1%	1%
Black	83%	36%	20%	16%	32%
White	18%	64%	80%	84%	68%
Total Membership	4505	6184	10205	3424	7440
Gifted	2%	7%	7%	6%	6%
Handicapped	11%	12%	12%	13%	10%
Reduced/Free Lunch	66%	24%	30%	23%	33%
Parents Less than H. S. Diploma	19%	14%	14%	18%	14%
Families Below Poverty Level	26%	15%	8%	15%	12%
Median Household Income	$10542	$15172	$16031	$15655	$14782

DIAL-R Systems

The DIAL-R School System 1 was a city school district located in the south-central section of North Carolina. Although this city school system had merged with the neighboring county system in 1993, the schools that made up the new merged system had remained intact. Because the system had not redistricted, the sixth-grade population at the sample school was the same group that had entered the city school district in 1988 except for changes which had occurred due to student mobility. The sixth-grade student population in this school was approximately 423.

DIAL-R School System 2 was also a city school district geographically located in the piedmont section of North Carolina in the same county as the PSB school system. Both systems were funded by the same county government. The system consisted of six elementary schools and two middle schools. The sixth-grade population in this system was approximately 476. All elementary school districts were represented in the sample. The DIAL-R sample consisted of a minimum of 100 students recruited from both school systems.

Gesell Systems

GSRT School System 1 was a city school district located in the central-piedmont section of the state. The school district consisted of five elementary schools, and two middle schools with an approximate sixth-grade enrollment of 154 students. All elementary school districts were represented in this study.

GSRT School System 2 was located in the center of the state and consisted of six elementary schools, and two middle schools with an approximate six-grade population of 607. Subjects were selected from students who were administered the GSRT after being screened by the DIAL (the forerunner of the DIAL-R) during preschool screening. The GSRT was administered to students falling in the at-risk category on the DIAL. All elementary schools were represented in the pool of subjects from this system. The parents of all students who received GSRT screening were asked for approval for their child's records to be included in the research. A target sample of 100 GSRT subjects was recruited from both school systems.

Preschool Screening Battery System

The PSB School System consisted of 12 elementary schools and 4 middle schools, and had a sixth-grade population of 786 students. A nonrandomized sampling procedure was employed in this district because preschool screening data were available for sixth grade students from only four of the twelve elementary schools. This situation occurred due to lack of consistent system-wide policies regarding storage and disposal of test data. The sixth-grade population from these four elementary schools was considered to be the accessible population for the PSB sample (Ary, Jacobs & Razavieh, 1972).

The PSB sample was considered to be representative of the total district population as each school represented a different geographic sector of the school system, and all geographic sectors were represented in the sample.

One school was located in a college town in the western section of the county, and the parents represented a diverse population from college professors and other professionals to lower income families. One school was located in the southwestern section of the county near Chapel Hill, North Carolina which is the home of the University of North Carolina at Chapel Hill. Residents of Chapel Hill have migrated into this area seeking a rural setting while bringing the influences of a university town with them.

The third school was centrally located, and included a cross-section of county residents from farmers and mill workers to professionals. The fourth school was located in the northeastern fringes of the county, and included a large proportion of minority students.

Examination of the socioeconomic factors, ethnicity, and characteristics of the student enrollment in the sample schools indicated that these schools were representative of the total sixth-grade population of the PSB school system, and that delayed purging of preschool screening data had occurred randomly. See Table 2 for a comparison between PSB school system characteristics and the PSB sample characteristics. The PSB sample consisted of a minimum of 100 subjects recruited from the PSB school system.

Table 2

Comparison of PBS System Characteristics and PBS Sample

	American Indian	Asian	Hispanic	Afro American	W
PSB System	1%	1%	1%	17%	80%
PSB Sample	0%	1%	2%	16%	81%
	Gifted	Handi-capped	Reduced/ Free Lunch	Total Membership	
PSB System	7%	12%	30%	10205	
PSB Sample	8%	13%	31%	1866	

A composite of system characteristics of all school systems selected for the study closely resembled the state averages reported for ethnicity, gifted and handicapped membership, and medium income. See Table 3 for a comparison of these figures.

Table 3

Comparison of State Averages and Research Sample

	American Indian	Asian	Hispanic	Afro American	W
State Average	1%	1%	1%	32%	66%
Research Sample	1%	1.2%	1%	37%	63%
	Gifted	Handi-capped	Reduced/ Free Lunch	Medium Income	
State Average	6%	12%	39%	$13265	
Research Sample	5.6%	11.6%	35.2%	$14436	

Instrumentation

Gesell School Readiness Test

The GSRT consists of these five subtests: (1) Paper and Pencil Test, (2) Cube Tests, (3) Incomplete Man, (4) Naming Animals Interests, and (5) Interview. GSRT items are primarily perceptual-motor and responses are obtained through clinical observations of the child by the evaluator (Kaufman, 1971). Analysis of the GSRT yields a developmental age which is considered to be a measure of "behavioral maturity" (Ilg & Ames, 1965).

According to Bradley (1985), one of the strengths of the GSRT is the inclusion of informative descriptions of the children's performance on each test. The GSRT also provides useful information about school-age children in several content areas related to readiness and school performance (Waters, 1985). Norms are available from age 3 to 9 years. Scoring tables present the percentage of children "at each normed age" (Waters, 1985, p. 610) who give specific responses. Specific methods for combining subtest scores to obtain an overall score are not included, and some degree of clinical judgement is necessary for scoring. Because of this subjectiveness, the Gesell Institute offers extensive training for examiners, and will not allow a school system to purchase their materials unless trained examiners are available to administer the test.

Developmental Indicators for the Assessment of Learning (DIAL-R)

The DIAL-R, a multi-dimensional norm-referenced test for use with children 2 to 6 years of age, consists of three subtests: (1) Motor, (2) Concepts, and (3) Language. Eight specific skills are assessed within these areas. The Gross Motor Area assesses the ability to throw, catch, jump, hop, skip, stand still, and balance. The Fine Motor area assesses the child's ability to match designs, build block designs, cut with scissors, copy geometric designs, copy letters, repeat a hand-clapping rhythm, and execute finger-touching commands (Johnson & Goldman, 1990).

The Concepts subtest assesses the child's ability to name colors, identify body parts, count, name letters, sort chips, understand position words and concepts such as big and little (Vacc et al., 1987). Language development is assessed by articulation, repeating numbers and sentence series, naming nouns and verbs, classifying foods, story-telling and problem solving. The DIAL-R is recommended for developmental screening and compares favorably with other screening instruments (Poteat, 1989).

Preschool Screening Battery (PSB)

The PSB consists of the Preschool Screening Scale (PSS), the Peabody Picture Vocabulary Test-Revised (PPVT-R), the Developmental Test of Visual Motor Integration, and the Expressive One Word Picture Vocabulary Test.

Preschool Screening Scale. The Preschool Screening Scale evolved over an eleven-year period and was revised several times. It includes the following subtests: (1) Fine Motor, (2), Visual Skills, (3) Auditory Skills, (4) General Information, and (5) Concepts. The form used in the Spring of 1988 included the following tasks: block building, visual sequencing, matching shapes, letters, and numbers, repetition of numbers and sentences, and following directions. Verbal tasks included answering general information questions and using concepts such as prepositions and numbers. Correct responses are summed to obtain subtest and total test scores. The PSS was used in conjunction with three other nationally normed tests to determine at-risk status.

Peabody Picture Vocabulary Test-Revised. The PPVT-R, a nonverbal multiple-choice test designed to measure auditory-receptive vocabulary, consists of 350 items or plates which have been compiled into two forms (Sattler, 1992). Each plate contains four pictures. The child is asked to point to the picture of the word spoken by the evaluator.

The PPVT-R was standardized on a sample of 4,200 subjects aged 2-5 to 18-0 which was representative of the U. S. population (McCallum, 1985). Standard scores, age equivalents, and percentiles are provided. Minorities were represented in the normative sample and in the test plates (McCallum, 1985). According to McCallum (1985), gender and ethnic stereotyping was eliminated in the revised version.

Developmental Test of Visual Motor-Integration. The DTVMI was designed to measure perceptual-motor ability for children aged 4 through 13 years (Beery, 1989). The standardization group for the 1982 norms consisted of 3,090 children from 2-9 to 19-8, and included subjects from various ethnic and socioeconomic groups (Sattler, 1992). Both sexes were represented equally (Sattler, 1992).

The test consists of 24 geometric shapes the subject is required to copy. Items are arranged in ascending order of difficulty. Raw scores are converted into standard scores which also yield percentile bands and age-equivalent scores (Sattler, 1992).

Expressive One-Word Picture Vocabulary Test. The EOWPVT assesses expressive language by measuring the child's ability to make associations between a picture of an object or a pictorial event. The examiner shows a picture to the subject, and then asks the subject to name the object or event shown in the picture. Norms are available for ages 2-0 to 11-11.

The California Achievement Test

The California Achievement Test (CAT) Form E was designed to measure performance in the basic skills of reading, language, and mathematics. This test battery, which was renormed in 1985 is a norm-referenced, nationally standardized achievement test. According to Airasian (1989) who reviewed this test for the Tenth Mental Measurements Yearbook, the CAT has been "a well-respected test battery for over 50 years" (p. 126).

Form E was normed on 300,000 students in grades K through grade 12 in Fall 1984 and Spring 1985. When compared with the U.S. 1983 Census Data, the norming sample was considered to be representative of the national school population (Airasian, 1989). Seven types of scores are available for the CAT E: scale scores, percentile ranks, normal curve equivalents, stanines, grade equivalents, anticipated achievement scores, and objective performance indices. Scale scores are considered to be the most technically desirable scores provided by the CAT Form E (Wardrop, 1989). Wardrop (1989) concurred with Airasian (1989) when he described the CAT is "one of the best standardized achievement batteries available" (p. 133).

North Carolina End-Of-Grade Reading Test

The North Carolina End-Of-Grade Reading Test yields a developmental rate for reading derived from student performances on a multiple choice reading test. The one hundred point developmental scale for reading was developed by the L. L. Thurstone Psychometric Laboratory at the University of North Carolina at Chapel Hill. Students are compared with other students at their grade level.

Standardization of End-of-Grade Testing began in 1993. Developmental scales were computed to measure students' progress in academic areas as they progress through school. Students are expected to obtain higher scores each year (Pommerich et al., 1993). Each developmental score is associated with a performance level. Level I indicates insufficient mastery for success at the next grade level. Level II represents inconsistent mastery and minimal preparation for the next grade level. Level III indicates consistent mastery of grade level subject matter, and Level IV represents consistently superior proficiency at grade level work. Student scores may be compared with those of the school, the school district, and the state (Pommerich et. al., 1993).

Data Collection

Informed consent forms, and introductory letters (Appendix B) were sent to parents of sixth-grade students in the target schools. Parents were asked to return signed consent forms to the classroom teacher within five school days. After two weeks, second letters and informed consent forms were sent to parents who did not respond to the first distribution. Novelty pencils were used as incentives to students for whom signed consent forms were returned.

Data were collected on each subject whose parents signed and returned the consent form giving approval. Subjects with preschool screening data were included in the data analysis. Demographic data were collected on gender, parental occupation, grade retentions, and placement in exceptional children programs. Socioeconomic status was determined by the parental occupation reported on consent forms. Duncan's Socioeconomic Status Indexes (SEI) were assigned to both parents. The parental occupation with the highest SEI score was entered in the data base. The updated TSEI version of the Duncan SEI was used. The original SEI is a standardized measure of the status of individual occupations which was based on open-ended questions, and existing data, and provides scores for all occupations used in the 1970 Census (Reiss, 1961; Stricker, 1988). The TSEI was updated to include occupations used in the 1980 Census.

Preschool scores for students with approved consent forms were collected from student cumulative records in the five school systems. Subtest scores and total test scores from the GSRT, the DIAL-R and the PSB were recorded.

The following subtest scores from the third-grade administration of the California Achievement Test were recorded: total reading, vocabulary, and word analysis. End-of-Grade Reading scores administered at the end of fourth and fifth grades were also collected. All test scores were obtained from student record reviews, and all reading grades entered in the cumulative folders were collected.

Data Analysis

Different multivariate analysis methods were computed to answer the research questions outlined this study. The predictive validity of the three screening batteries was assessed by performing predictive discriminant analysis to establish procedures for classifying subjects into reading groups based of preschool screening scores and demographic variables. Third grade membership in the three reading groups was determined by total reading subtest scores on the California Achievement Test and reading grades at the third grade level. Reading group membership at the fifth grade was determined by the North Carolina End-of-Grade Reading scores and fourth and fifth level reading grades. These two measures for each grade level were converted into z-scores and then averaged into a single variable. Standard deviation cutoffs were used to form the following groups (Fletcher at al., 1984).

Below Average	-.5 s. d. below the mean
Average	-.4 s. d. below the mean to +.4 above the mean
Above Average	+.5 s. d. above the mean

Principal component analysis and principal axis factoring procedures were computed to reduce the number of predictor variables in the Gesell and the Preschool Screening Battery analyses, and to identify underlying constructs. Predictor variables were combined into surrogate variables in the Preschool Screening Battery and the Gesell analyses.

Predictor variables from the preschool screening batteries and demographic variables were entered into predictive discriminant analysis. This statistical procedure was selected because it is useful when the goal of analysis is to find the dimensions along which groups differ, and to find classification functions or constructs to predict group membership (Hair et al., 1992; Steven, 1992).

Bartlett's Chi Square Test was used to determine total significance, and to determine the number of significant discriminant functions (Stevens, 1992). The discriminant procedure used in this analysis was designed to develop the "best one-variable model, followed by the best two-variable model and so forth, until no other variables meet the desired selection rule" (Hair, Anderson, Tatham & Black, 1992, p. 111). The selection rule used in the procedure maximizes the distance between groups (Hair et al, 1992).

Canonical discriminant functions or linear composites were calculated from this three-group discriminant model. In this procedure, the first function is extracted and the chi-square is recalculated. If significant differences are present in the remaining variance, additional discriminant functions are computed. This procedure is repeated until the maximum number of functions are extracted (Hair et al., 1992). The number of possible discriminant functions is (k-

1) in which k = the number of groups (Stevens, 1992, p. 275).

Because it is desirable to test how well the classification coefficients from the research sample generalize to other samples, cross-validation was obtained by randomly dividing the total sample into two parts (Hair et al., 1992). The original sample was divided into two groups (analysis and holdout group) in a 50 - 50 split. The holdout sample was selected through a proportional, stratified, sampling procedure. The sizes of the groups were proportionate to the sample distribution.

The analysis sample was used to compute the discriminant functions, and the coefficients from that computation were used to obtain discriminant scores for the holdout sample (Hair et al., 1992). Classification matrices were developed for the two groups. Hit ratios (percentage of correctly classified cases) were computed to determine how well the discriminant functions "classified the statistical units" (Hair et al., 1992, p. 100). Classification accuracy was determined by comparing the hit ratios from the classification matrices with the proportional chance criterion which was selected because it is recommended when group sizes are unequal, and when the researcher is interested in classifying members into two or more groups (Hair et al., 1992). The formula used to compute the proportional chance criteria is listed below (Huberty, 1994).

$$e = \sum_{g=1}^{k} q_g \, n_g$$

q = estimated prior probability

ng= number in group g

The classification accuracy of the discriminant functions must be at least 25 percent greater than the proportional chance criterion in order to be considered significant. This criterion is suggested by Hair et al. (1992) as a guideline for determining significant differences between classification accuracies.

Press's Q statistic and the standard normal statistic were used to test the discriminatory power of the classification matrices and individual groups. The Press's Q statistic was used to compare the number of correct classifications with total sample size and the number of groups. The computed value was compared with the chi-square value for 1 degree of freedom at 0.01 confidence level. The formula for Press's Q statistic is below (Hair, 1992. p. 106).

$$\text{Press's } Q = \frac{\{N - (n \times K)\}^2}{NK(-\ 1)}$$

where

N = Total sample size
n = Number of observations correctly classified
K = Number of groups

A standard normal statistic was used to test the null hypothesis that the observed classification accuracy was not different than that expected by chance alone. The following formula was used to compute z-scores for overall group results (Huberty, 1994, p. 105).

$$Z = \frac{o - e}{\sqrt{e(N - e)/N}}$$

e = proportional chance criteria

N = total number in group

Standardized normal statistics were also computed for separate groups using this formula (Huberty, 1994, p. 105).

$$Z = \frac{nhits - eg}{\sqrt{eg\ (ng - eg/ng}}$$

nhits = number of correctly classified subjects

ng = number in group g

An effect size statistic was computed to determine if the results from the classification matrices were actually better than those which would have been derived by chance

(Huberty, 1994). This statistic can be interpreted to measure the "meaningfulness" of the effect size (Huberty, 1994).

The following formula was used (Huberty, 1994).

$$I = \frac{Ho - He}{1 - He}$$

Ho = observed hit rate

He = hit rate expected by chance

The discriminant functions were interpreted in an attempt to explain the relationship between the dependent categorical variables or group membership and predictor variables (Hair et al., 1992). The following factors were expected to be identified as underlying constructs of the screening batteries: a language/comprehension factor, a perceptual/perceptual-motor factor, and a conceptual factor (Fletcher et al., 1984; Kaufman, 1971; Roth et al., 1993; Stanovich et al., 1984; Tramontana et al., 1988; and Vacc et al., 1987).

Discriminant functions were interpreted through standardized coefficients, discriminant function-variable correlations, also called discriminant loadings, and partial F values (Hair et al., 1992; Stevens, 1992). Discriminant coefficients were used to check for redundancy. Because they are similar to beta weights in regression analysis, a small discriminant coefficient may indicate its' corresponding variable is irrelevant, or that it shares a high degree of multicollinearity with other variables (Hair et al., 1992).

Discriminant loadings were used to name the underlying constructs represented by the discriminant functions, because the loadings depict the variance the independent variables share with the discriminant function (Stevens, 1992). The partial F values, derived from the stepwise computational approach, were ranked, and larger F values were interpreted as having greater discriminating power (Hair et al., 1992). The Verimax rotation procedure in which the variance is redistributed was employed to aid in substantive interpretation of the constructs (Huberty, 1994).

The effects of the predictor variables on reading achievement and specific reading subskills were analyzed through supplementary multiple regression and partial correlational analyses. The influence of the predictor variables on reading acquisition was studied through interpretation of the discriminant functions. Additionally, the influence of predictor variables on specific reading subskills was assessed through the supplementary regression analyses.

Finally, students who were screened as "ok" at preschool who developed reading problems later, and those who were predicted to be at-risk by preschool screening who actually became adequate readers were studied. Analyses were conducted on those groups to determine the significance of differences between predictor mean scores of subjects with inaccurate classifications and those who were correctly classified.

Summary

In conclusion, this study focused on the predictive validity of predictor variables measured on three preschool screening batteries to predict reading group membership at the end of elementary school. The relationship of underlying constructs extracted from the batteries and reading subskills was also studied. Methodological procedures used to answer the research questions derived from these issues were factor analysis, predictive discriminant analysis, and regression analysis.

CHAPTER IV

Results of the Study

The results of the data analysis described in Chapter III will be presented in this chapter. The process of data collection is discussed in the first section. Procedures employed in the statistical analysis are presented in the second section. The results of the analysis are presented in the third section and a summary concludes the chapter.

Data and Subjects

Data were collected through a two-step process. First, introductory letters and informed consent forms were sent home to parents or legal guardians. Two weeks later follow-up letters and informed consent forms were sent to parents who had not responded to the first letter. Response rates and percentages for each preschool screening battery are presented in Table 4.

The total number of possible subjects across groups was 1051, and the final return rate for all groups combined was 54% (N=575). This rate was partitioned into three groups, <u>no</u> responses (parents refused permission), <u>yes</u> responses (parents gave permission) and data were available, and <u>yes</u> responses but data were not available. Twenty-nine percent of the <u>yes</u> responses were eliminated due to the absence of preschool screening data in the student's cumulative folder.

Table 4

Informed Consent Response Rates

DIAL-R		N	%
Group	Total returned	234	50.2
	Yes	184	39.5
	No	51	10.9
	Yes with data	101	21.7
GSRT	Total returned	153	50.5
	Yes	119	39.3
	No	24	7.9
	Yes with data	95	31.4
PSB	Total returned	188	66.7
	Yes	154	54.7
	No	34	12.1
	Yes with data	103	36.8
Total	Total returned	575	54.7
	Yes	466	44.3
	No	109	10.4
	Yes with data	298	28.4

Note. There were 466 possible DIAL-R subjects, 303 GSRT subjects, 282 PSB subjects, and 1051 in the total group.

The final number of potential subjects was 298 (101 in the DIAL-R group, 94 in the Gesell group, and 102 in the Preschool Screening Battery group). Two Gesell subjects were dropped due to incomplete data. The final sample consisted of 296 subjects, 147 males (49%) and 151 females (51%). The ethnic makeup of the group was 103 white males and 104 white females (69%) and 42 non-white males and 46 non-white females (32%).

Statistical Analysis

Predictive discriminant function analysis was computed to assess the predictive validity of three preschool screening batteries, the DIAL-R, the Gesell School Readiness Test, and the Preschool Screening Battery. The purpose of the analysis was to determine if predictor variables from preschool screening batteries were predictive of later reading achievement. SPSS for Windows Advanced Version V6.1 was used for the analysis (SPSS, 1995).

Grouping Variables

Grouping variables were formed by combining reading achievement scores and reading grades. At the third grade, Total Reading scores from the California Achievement Test were combined with Grade 3 reading grades. Both variables were converted to z-scores before being averaged into one grouping variable, Reading 3. Scale scores from the California Achievement Test were used for the computation and reading grades were converted to numerical scores using the following 4-point scale.

A =4.0	A- =3.7	B+ =3.5	B =3.0	B- =2.7	C+ =2.5
C =2.0	C- =1.7	D+ =1.5	D =1.0	D- = .7	F =0

The following standard deviation cutoffs were used to form the reading groups (see Fletcher & Satz, 1980).

Below average	-.5 s.d. below the mean
Average readers	-.4 s.d. below the mean to +.4 above the mean
Above average	.5 s.d. above the mean

End of Grade Reading scores and reading grades from Grade 4 and Grade 5 were combined into the Reading 5 grouping variable. Because two school systems did not assign Grade 5 reading grades, grades from the end of the Grade 4 were used as the reading average for students in those schools. Modified grades (those with notations in the cumulative folders indicating they were modified due to placement in exceptional children classes) were marked down one grade level to adjust for more lenient grading procedures in special classes. Grades were converted using the 4-point scale. Predictors were converted to z-scores before being averaged into one grouping variable (Reading 5). Groups were coded as 1 = below average readers, 2 = average readers, and 3 = above average readers. Due to low numbers in the smaller groups, it was not feasible to use 1 standard deviation below the mean as the cutoff score for severely disabled readers as originally planned. Consequently, the severely disabled readers and the mildly disabled readers group were combined into a single below average group.

Preschool Screening Variables

Data reduction techniques were employed to reduce the number of predictor variables and to identify any latent constructs in the original variables (Hair et al., 1992; Huberty, 1994). Factor analyses were conducted using discriminant analysis and principal axis factoring procedures of the Data Reduction option of SPSS for Windows Advanced Version, V6.1 (SPSS, 1995). The goal of data reduction was to obtain estimators with lower bias and higher precision (Huberty, 1994). Factor analysis of the DIAL-R was not conducted because raw scores for individual items were not available.

Factor Analysis

Gesell School Readiness Test (GSRT). The developmental age scores from the Gesell subtests were converted to months before being entered into the data file. Subtest scores were entered into principal axis factoring analysis, and an a priori criterion for extracting three factors was set in an effort to replicate Kaufman's (1971) previous analysis. Kaufman's factors were:

Factor I - Paper and pencil coordination on which Copy Forms loaded heavily

Factor II - Awareness of part-whole relationships on which the Incomplete Man subtest loaded heavily

Factor III - Academic achievement on which Writing Name and Numbers loaded heavily (Kaufman, 1971, p. 1355).

In the Gesell analysis, the latent root criterion was used to determine the number of significant factors along with the percentage of variance criterion, and the screen test criterion (Hair et al., 1992, p. 237). In deriving a screen test, the latent roots are plotted against the optimum number of factors (Hair et al., 1992). According to Hair et al. (1992) "the point at which the curve first begins to straighten out is considered to indicate the maximum number of factors to extract" (p. 238).

Using the latent root criterion, one significant factor was extracted with an eigenvalue of 3.73. However, the screen plot test indicated that two factors were significant. The factor solution with two factors accounted for 66.2% of the cumulative variance. According to Hair et al. (1992), it is not uncommon in the social sciences to consider a solution that accounts for 60 percent to be satisfactory (p. 237).

A second principal axis factoring was conducted in which two factors were extracted accounting for 66.2% of the variance. Again, one factor was significant using the latent root criterion (Eigenvalue, 3.72), and the screen test indicated two factors were significant. Based on this data, the variable loadings on the two factor solution were interpreted.

The Incomplete Man and other perceptual motor tasks (Copy Forms and Cubes) loaded heavily on the first factor. This pattern is reminiscent of the factor Kaufman (1971) identified as behavioral maturity.

The Interview subtest also contributed to the common variance of the first factor. Although this subtest does not consist of perceptual motor tasks, it does require knowledge or awareness of self and family, attributes which might be considered to be a component of social maturity.

Factor II was similar to Kaufman's Factor III which was named Achievement. The subtests, Writing Names and Numbers and Naming Animals and Interests loaded on Factor II. Although the writing tasks required paper and pencil coordination, they were also considered to be dependent on training and experience (Kaufman, 1971, p. 1355). Naming Animals and Interests, which are verbally loaded tasks are also likely to be dependent on training and exposure to relevant experiences. Table 5 for factor loadings.

Table 5

Gesell Factor Loadings From Principal Axis Factoring

Factor I		Factor II	
Incomplete Man	0.87	Writing Name and Numbers	0.88
Copy Forms	0.68	Naming Animals	0.53
Interview	0.55		
Cubes	0.45		

Due to limited reliability data on GSRT subtests, surrogate variables were formed (Hair et al., 1992; Waters, 1985). Gesell subtest scores were converted to z-scores then averaged into two variables, Maturity and Achievement. The four subtests which loaded on Factor I, were combined into the GSRT Maturity variable. The Writing Name and Numbers and Naming Animals and Interests subtests were combined into the GSRT Achievement variable. Reliability coefficients of 0.84 and 0.80 were computed for GSRT Maturity and GSRT Achievement respectively.

Preschool Screening Battery (PSB). Factor analyses were computed with the nine subtests of the Preschool Screening Battery. Subtest scores were entered into the data file as raw score totals. Exploratory principal axis factoring and principal component analyses were conducted to reduce the number of PSB variables. The latent root criterion was used to determine significance and eigenvalues of 1 were considered significant. Three of the factors had eigenvalues of 1 or greater. The fourth factor with a heavy Colors loading had an eigenvalue of 0.95. (The Colors subtest consistently loaded on a separate factor in principal axis factoring and principal component analyses.) The Verimax rotation procedure was used to obtain orthogonal solutions, and the Rotated Factor Matrix was used to interpret the factors. Results are listed in Table 6.

Based on results of factor analysis and theoretical considerations, the Numbers, Blocks, General Information, and Readiness subtests were combined into one variable (Cognition). This variable was named Cognition because it was comprised of a variety of subskills related to conceptual knowledge.

Table 6

PSB Factors Extracted From Principal Components Analysis

Factor 1		Factor 2		Factor 3		Factor 4	
Blocks	0.82	Aud Mem	0.56	Commissions	0.84	Colors	0.97
Numbers	0.81	Vis Mem	0.70				
Gen Infor	0.74	Vis Dis	0.79				
Readiness	0.54						

In order to enhance predictive accuracy, Factors 2 and 3 were collapsed into one variable, Perception. These variables were combined under the premise that Auditory Memory (repetition of numbers and sentences), Commissions (following one, two, and three-part directions), Visual Memory (identification of missing items), and Visual Discrimination (matching) were all types of perceptual tasks. Reliability coefficients of 0.74 and 0.55 were obtained for Cognition and Perception, respectively.

The Colors subtest was included as a separate variable due to exploratory predictive principal component analyses. In these analyses, the Colors subtest significantly improved the classification accuracy for low and average achieving readers at Grade 5.

Developmental ages for the PPVT-R and EOWPVT were converted into months before being entered into the data file. These variables were converted to z-scores and combined into one language variable (PSB Language). A reliability coefficient of 0.778 was obtained for the combined variable.

In all cases, original variables were converted to z-scores and then averaged into surrogate variables (Hair et al., 1992). The surrogate variables were then entered into predictive discriminant analyses for classification purposes.

Demographic Variables

Socioeconomic status was determined by parental report of occupations on the consent form. Occupations reported were matched as closely as possible to an occupational class on the Duncan Socioeconomic Index (Stevens & Cho, 1985). Socioeconomic Index (SEI) scores were determined for both parents. For two parent families, the highest of the two SEI scores was entered in the data file.

Other demographic variables were gender and ethnicity. Gender was coded as male = 0 and female = 1. Ethnicity was coded as 1 = white and 2 = non-white. See Table 7 for a final listing of predictor variables, demographic variables, and grouping variables used in the six discriminant analyses.

Discriminant Function Analysis

Predictor variables were entered into discriminant analyses to determine their power to predict reading group membership at Grade 3 and Grade 5 and to identify underlying constructs. Demographic variables (SES, Ethnicity, and Gender) were included in second analyses to control for their effects on reading achievement. Each analysis was computed with third grade data and again with fifth grade data. When predictor variables were entered into the analyses, one significant function accounted for 75 to 98%

Table 7

Variables Used in Predictive Discriminant Analysis

	Combined Variables
DIAL-R	
1. Motor	Original subtest
2. Concepts	Original subtest
3. Language	Original subtest
GSRT	
4. Maturity	Cubes, Incomplete Man, Copy Forms
5. Achievement	Writing Names/Numbers, Naming Animals/Interests
PSB	
6. Perception	Aud. Memory, Vis. Discrim./Memory, Commissions
7. Cognition	Blocks, General Information, Readiness, Numbers
8. Language	PPVT, EOWPVT
9. Colors	
Demographic Variables	
10. SES	Duncan's SES Index
11. Gender	0=Male, 1=Female
12. Ethnicity	1=White, 2=Non-white
Grouping Variables	
13. Group 3	Tot. reading score (CAT), reading grade
14. Group 5	4th & 5th EOG Reading scores, reading grade ave.

of the variance in all cases. Fletcher & Satz (1980) reported similar findings when studying the developmental changes in reading achievement. They found 90% of the variance in reading was explained by one composite. Percentages of variance per function are listed in Table 8.

Table 8

Percentage of Variance Explained by Function 1 for DIAL-R, GSRT and PSB

	Third Grade		Fifth Grade	
	P Value	% of Variance	P Value	% of Variance
DIAL-R	0.0000	98%	0.0219	75%
GSRT	0.0038	89%	0.0068	99%
PSB	0.0026	93%	0.0001	94%

Factors with eigenvalues of 1 or greater were retained. With the exception of preliminary and exploratory analysis, the direct method (all variables entered at the same time) was used as recommended by Huberty (1994). The linear classification rule was used due to the small to moderate n:p ratios where n is the number of subjects per group and p is the number of variables (Huberty, 1994, p. 64). Hit rate estimators were based on group sizes of the holdout sample.

Group Probability Estimates. Unequal prior probability estimates were used in the analyses due expected unequal probabilities of group membership in the sample population

(Huberty, 1994). This expectation was based on state frequency distributions in reading reported for the 1988 administration of the California Achievement Test. According to the report, 25% of the third grade students scored in the 7th, 8th, and 9th stanines (Well Above Average range). The Slightly Above Average and Average groups scored in the 5th and 6th stanines representing 38% of the students. The Slightly Below Average and Well Below Average groups scored in the 1st, 2nd, 3rd, and 4th stanines which represented 26% of the students (Moore & Zuckerman, 1992).

Expectations associated with performance levels reported for the North Carolina End of Grade testing were: 39% of students at Levels I and II (demonstrating inconsistent or insufficient mastery of knowledge and skills), 40% of students at Achievement Level III (demonstrating mastery of grade level subject matter), and 20% of students at Level IV (performing in a superior manner) on the End of Grade Reading test (North Carolina Department of Public Instruction, 1994). Because the sample was considered to be representative of sixth grade students in North Carolina, prior probabilities were based on sample proportions (Hair et al. 1992, p. 111).

Validation. External classification analysis was conducted through cross validation. The holdout method was used which utilizes a classification rule determined by an analysis group and validated on the holdout group (Huberty, 1994, p. 87). The classification rates reported were obtained from the validation group. The SPSS for Windows, Advanced Version V6.1 Classify procedure was used to compute the analysis (SPSS, 1995).

A proportionately stratified sampling procedure was used to select subjects for the analysis group (Hair et al., 1992). The total group for each preschool screening battery was divided into two groups representing a 50-50 split. The analysis and holdout groups were matched on separate group sizes, gender, ethnicity, and socioeconomic status.

Multivariate Normality and Equality of Variance. Multivariate normality was determined through review of stem and leaf plots and normal probability plots computed by the Statistics function (Exploring Data) of SPSS for Windows Advanced Version V6.1 (SPSS, 1995). The Box M statistic was computed to check for group covariance homogeneity (Huberty, 1994). The condition of homogeneity was met for the GSRT and PSB analyses. It was not met for the third grade DIAL-R analysis (without and with demographics) which resulted in F transformations of the Box M statistic of 2.378 ($P = 0.0047$) and 2.333 ($P = 0.0056$). The DIAL-R variables (Concepts, Motor, and Language) were transformed by the Transformation procedure (Rank Cases option) of SPSS for Windows Advanced Version V6.1 (Norusis, 1993). The ranked DIAL-R variables were entered into predictive discriminant analysis for the third grade which yielded an insignificant Box M statistic (0.88532, P= 0.5614). Demographic variables were then added resulting in an Approximate F value of 1.00974, (P.= 0.4537) indicating group covariance homogeneity for both DIAL-R analyses was met.

The rank transformation was used because it is applicable to non-normal continuous distributions according to Huberty (1994). In this procedure, all the observations for a specific predictor are "pooled and then ranked from 1 for the smallest observation to N for the largest" (Huberty, 1994, p. 148).

Results

Total group results and separate group results of the six predictive discriminant analyses which included the predictor variables and the grouping variables at third and fifth grade will be discussed in detail. Additionally, Table 9 shows the total group results from six classification matrices in which the hit rates represent the percentage of subjects correctly classified, and Press's Q statistic which represents the discriminatory power of the classification matrix when compared to the chance model (Hair et al., 1992). Press's Q statistic was compared to the critical value (6.63) from the chi square distribution for 1 degree of freedom at 0.01 confidence level. Results are reported for principal component analyses with and without demographic variables at the third and fifth grades.

Table 9

Classification Accuracy at the Third Grade With and Without Demographics

Third Grade				
Preschool Battery	Hit Rate		Press's Q	
	Demographics			
	Without	With	Without	With
DIAL-R	55.10%	61.22%	10.44*	17.15*
Gesell (GSRT)	43.48%	45.65%	2.13	3.14
Preschool Screen B.	62.75%	62.75%	19.85*	19.85*
Fifth Grade				
	Hit Rate		Press's Q	
	Demographics			
DIAL-R	59.18%	49.03%	14.73*	6.25
Gesell (GSRT)	44.44%	51.11%	2.50	6.40*
Preschool Screen B.	70.00%	58.82%	25.51*	5.21*

Relative Predictive Validity of the Preschool Screening Batteries

Overall results from the Preschool Screening Battery were better than chance according to Press' Q statistic at both grade levels with and without demographics. DIAL-R results were significant at the third grade level with and without demographics and at the fifth grade level without demographics. Gesell classification results were significant at the fifth grade level with demographics included. Caution should be used in interpreting the results of Gesell and Preschool Screening Battery classification results (with demographics) due to small n/p ratios. In order to estimate true hit rates, Huberty (1994) proposed that the number of subjects in the smallest group be at least five times the number of predictors where n1 is equal to or greater than 5p (p. 97). The predictors were increased from 2 to 5 for the Gesell and from 3 to 6 for the PSB when demographic variables were included. The number of predictors to total group membership of the smallest group exceeds the recommended ratio for the PSB at Grade 3 (n1=16, p=6) and at Grade 5 (n1=24, p=6) and the GSRT at Grade 5 (n1=24, p=5). Because demographics were included in the analyses to control for their effects rather than to determine their discriminatory power, the results of the classification matrices with demographics were used solely for comparison purposes in this research.

Discriminant functions derived from the GSRT and PSB analyses (with demographics) are not suggested for use in future predictive discriminant computations. The DIAL-R did have adequate n/p ratios to determine true hit ratios, and the inclusion of demographics at the third grade level improved the classification power of the DIAL-R somewhat. The inclusion of demographics in the DIAL-R third grade analysis resulted in correct classifications of two additional Group 3 subjects. Vacc et al. (1987) reported similar results when demographics were entered into canonical correlational analysis of DIAL variables and the California Achievement Test scores at the first grade. These researchers found SES, gender, race, age, and parental status contributed little to overall reading scores.

At the fifth grade, inclusion of demographics decreased the discriminatory power of the DIAL-R to below the chance level. The major difference in the fifth grade analyses with and without demographics was in the discriminatory power of the predictor variables to classify subjects in the Group 2 category. The hit rate for Group 2 (without demographics was 55.6%. In contrast, when demographics were added to the analysis, the hit rate for Group 2 subjects decreased to 16.7% (3 out of 18). This decrease in correct classifications for Group 2 subjects resulted in a significantly lower overall hit rate affecting the Press's Q statistic in which correct classifications are compared with total sample size.

Classification Matrix Accuracy

DIAL-R. The Motor, Concepts, Language, and Group 3 variables were entered into Predictive Discriminant Analysis. The DIAL-R total group sample included 18 white males, 29 white females, 26 non-white males and 28 non-white females. The analysis group consisted of 9 white males, 15 white females, 13 non-white males, and 14 non-white males (n= 51). See Table 10 for classification results.

Table 10

DIAL-R Classification Matrices for 3rd Grade and 5th Grade

Actual Group	No. of Cases	Predicted Group		
		1	2	3
Third Grade				
1	20	18	1	
		90.0%	5.0%	5.0%
2	15	7	2	6
		46.7%	13.3%	40.0%
3	14	2	5	7
		14.3%	35.7%	50.0%
Fifth Grade				
1	16	9	5	2
		56.3%	31.3%	12.5%
2	18	5	10	3
		27.8%	55.6%	16.7%
3	15	0	5	10
		0.0%	33.3%	66.7%

Gesell. Two GSRT variables, Maturity and Achievement were entered into the analyses with the grouping variables, Group 3 and Group 5. Total group membership for the GSRT sample was 93 (44 white males, 24 white females, 12 non-white males and 11 non-white females). The analysis sample consisted of 21 white males, 13 white females, 7 non-white males and 5 non-white females. See complete results in Table 11.

Table 11

GSRT Classification Matrices at the 3rd Grade and 5th Grade

Actual Group	No. of Cases	Predicted Group		
		1	2	3
Third Grade				
1	14	4	6	4
		28.6%	42.9%	28.6%
2	16	4	6	
		25.0%	37.5%	37.5%
3	16	0	6	10
		0.0%	37.5%	62.5%
Fifth Grade				
1	12	3	6	3
		25.0%	50.0%	25.0%
2	18	2	11	5
		11.1%	61.1%	27.8%
3	15	1	8	6
		6.7%	53.3%	40.0%

Preschool Screening Battery. Three variables from the PSB were entered into predictive discriminatory analysis at the third grade level, Cognition, Perception, and PSB Language. The PSB sample consisted of 102 subjects (40 white males, 50 white females, 5 non-white females, and 7 non-white males). The analysis group consisted of 51 subjects, 20 white males, 25 white females, 3 non-white males and 3 non-white females.

The Colors variable was added to the PSB predictor set at Grade 5. The inclusion of the Colors variable in the fifth grade analysis improved the actual classification ratio for Group 1 from 16.70 % (without Colors) to 66.78% (with colors). The inclusion of the Colors variable did not improve the classification rate at the third grade, however. Likewise, the inclusion of the DTVMI did not improve classification results at either grade level. Reynolds et al. (1980) found that adding a verbal or nonverbal measure to an existing verbal or nonverbal measure when predicting reading achievement at the second grade improved prediction only slightly. Duffey et al. (1976) found that kindergarten DTVMI scores had little practical predictive validity when correlated with the Reading subtest of the Sanford Achievement Test administered at the end of second grade. Classification results are reported in Table 12.

Table 12

PSB Classification Matrices at the 3rd Grade and 5th Grade

Actual Group	No. of Cases	Predicted Group		
		1	2	3
Third Grade				
1	8	3	4	1
		37.5%	50.0%	12.5%
2	19	5	11	3
		26.3%	57.9%	15.8%
3	24	0	6	18
		0.0%	25.0%	75.0%
Fifth Grade				
1	12	8	4	0
		66.7%	33.3%	0.0%
2	17	6	8	3
		35.3%	47.1%	17.6%
3	22	0	4	18
		0.0%	18.2%	81.8%

Classification Accuracy Measured by Z-Scores

The effectiveness of the six discriminant function analyses of the predictor variables derived from the preschool screening batteries and the grouping variables at the third and fifth grades will be discussed. The effectiveness of separate group predictions will be reported in addition to overall group results.

A standard normal statistic was used to test the null hypothesis that the observed classification accuracy was not different than that expected by chance alone. Proportional chance criteria were computed by first estimating prior probability of membership for each group by dividing the actual number of subjects in each group by the total sample size. Next, the expectancies (e) for each group were computed by multiplying the number in each group (n) in the validation sample by the estimated prior probability for each group. The chance probabilities for each group were then summed to obtain the proportional chance criteria.

Standard normal statistics were also computed comparing the overall results of each predictive discriminant analysis with the maximum chance criterion by substituting the maximum chance criterion for the proportional chance criterion in the formula. The maximum chance criterion was computed by dividing the number of subjects in the largest group by the total number in the sample (Huberty, 1994).

The DIAL-R was better than chance at predicting total group membership at Grade 3 at the 0.01 level. The third grade predictions were better than chance for Groups 1 and 2 at the 0.05 level. Fifth grade DIAL-R results were better than chance for Groups 1 and 3 and for total group results at the 0.01 level. Group 2 results were better than chance at the 0.05 level.

When compared to the maximum chance criterion, the DIAL-R predicted total group membership was better than chance at the .01 level for the third and fifth grades. Complete results are listed in Table 13.

Table 13

Classification Results for the DIAL-R Based on Z-Scores

Group	Hits	n_g	e	z	P
Third Grade DIAL-R Holdout Group					
DIAL-R n1=(43) n2=(29) n3=(29) (101)					
1	18	20	6.60	5.42*	0.018
2	2	15	5.40	1.83*	0.034
3	7	14	4.34	1.51	0.066
Total Group	25	50	16.34	3.23*	0.002
Fifth Grade DIAL-R Holdout Group					
Total N DIAL-R n1=(33) n2=(36) n3=(31) N=101					
1	9	16	5.28	1.97*	0.000
2	10	18	6.48	1.72*	0.043
3	10	15	4.65	2.17*	0.000
Total Group	29	49	16.41	3.81*	0.000
Maximum Chance Z-Score 3.25, P. 0.000					

Note. e= proportional chance criterion, n_g = numbers in holdout groups.

Significant results were reported by the GSRT analysis for Group 3 at Grade 3 and for Group 2 at Grade 5. Overall group results were not significant at either grade level when compared to the proportional chance criterion or the maximum chance criteria. Table 14 contains z-scores associated with classification accuracy.

Table 14

Classification Results for GSRT Based on Z-scores

Group	Hits	n_g	e	z	P
Third Grade GSRT Holdout Analysis					
GSRT n1=(27) n2=(33) n3=(33) (93)					
1	4	14	4.06	0.040	0.399
2	6	16	5.68	0.167	0.433
3	10	16	5.68	2.19*	0.014
Total Group	20	46	15.42	1.43	0.076
Maximum Chance Z-score 1.25, P. = 0.106					
Fifth Grade GSRT Holdout Analysis					
GSRT n1=(24) n2=(35) n3=(32) N=91					
1	3	12	3.10	0.063	0.264
2	11	18	6.77	3.85*	0.000
3	6	15	5.16	0.514	0.305
Total Group	20	45	15.02	1.58	0.057
Maximum Chance Z-Score 0.38, P. 0.352					

Note. e= proportional chance criterion, n_g = numbers in holdout groups.

The third grade PSB analysis resulted in better than chance predictions for Groups 1 and 2 at the 0.05

probability level and at the 0.01 level for Group 3 and for total group results.

PSB fifth grade predictions were better than chance at the 0.01 level for Groups 1, 3, and total group results. Classification results for Group 2 ($z = 1.62$, $p = 0.053$) were not better than chance at the 0.05 level. When compared to the maximum chance criterion the PSB total group results were significantly better than chance at the 0.01 level for both grades. Complete results are reported in Table 15.

Effect Size

An effect size statistic was computed to determine if the results from the classification matrices were actually better than those which would have been derived by chance (Huberty, 1994). One way to interpret the I-value is in terms of reduction in error. For example, an I-value (0.53) may be interpreted as a reduction in classification errors of 53% I-indexes for the predictive discriminant analysis computations were included in Table 16. According to Cohen (cited by Sedlmeir & Gigerenzer, 1989) 0.18 = small effects, 0.48 = medium effects, and 0.83 = large effects. Using these ratings, I-values for the PSB were equal to medium size effects, and the I-value for the DIAL-R (0.39) approached medium size effects with the addition of demographics in the discriminant analysis at the third grade.

Table 15

Classification Results for PSB Based on Z-scores

Group	Hits	n_g	e	z	P
Third Grade PSB Holdout Analysis					
PSB n1=(16) n2=(38) n3=(48) (102)					
1	3	84	1.256	1.70*	0.044
2	11	19	7.087	1.88*	0.030
3	8	24	11.307	2.78*	0.003
Total Group	25	50	19.65	3.55*	0.000
Maximum Chance Z-score 2.21, P. = 0.000					
Fifth Grade PSB Holdout Analysis					
PSB n1=(24) n2=(35) n3=(43) N=102					
1	7	10	2.35	3.47*	0.000
2	9	17	5.83	1.62*	0.053
3	17	22	9.28	3.34*	0.000
Total Group	32	51	17.46	4.29*	0.000
Maximum Chance Z-Score 2.83, P. 0.013					

Note. e= proportional chance criterion, n_g = number in holdout group.

Table 16

Improvement Over Chance Indexes Third Grade and Fifth Grade

	Without Demographics	With Demographics
Effect Size Third Grade		
DIAL-R	I = 0.30	I = 0.39
GSRT	I = 0.14	I = 0.17
PSB	I = 0.47	I = 0.40
Effect Size Fifth Grade		
DIAL-R	I = 0.33	I = 0.25
GSRT	I = 0.17	I = 0.27
PSB	I = 0.53	I = 0.38

Interpretation of Functions

Functions extracted from the predictive discriminant function analyses were interpreted through examination of the discriminant loadings. Discriminant loadings from the structure matrices for Grades 3 and 5 are listed in Table 17.

The structure matrices and the group centroids were examined for each battery to determine if any early determinants of reading acquisition could be identified across batteries. The discriminant loadings and group centroids from third and fifth grades were also examined to detect any pattern among the constructs in their contribution to reading acquisition over time.

Table 17

Discriminant Loadings for the DIAL-R, GSRT, and PSB

	Third Grade		Fifth Grade	
	Function 1	Function 2	Function 1	Function 2
DIAL-R				
Concepts	0.88429	-0.21908	0.98148	-0.08489
Motor	0.57655	0.71855	0.78975	-0.25759
Language	0.70049	0.16203	0.74103	0.66776
Gesell				
Maturity	0.99947	0.03246	0.72521	0.68853
Achievement	0.78777	0.61597	0.14718	0.98911
Preschool Screening Battery				
Language	0.90353	-0.41981	0.94127	-0.33663
Cognitive	0.76050	0.57740	0.68759	0.03208
Perception	0.54527	0.48976	0.56563	0.53357
Colors	N/A	N/A	0.22746	0.70850

In the DIAL-R third grade analysis, the variables, Concepts and Language, had the highest loadings on Function 1. The Motor subtest had the highest loading on Function 2.

According to group centroids, Group 1 subjects had positive scores on the Motor subtest and negative scores on the Language and Concepts subtests. In contrast, Group 3 membership was influenced most by Concepts and Language. See listing of Group Centroids in Appendix G.

At the fifth grade, Concepts loaded highest on Function 1 followed by the Motor and Language subtests with loadings of 0.78975 and 0.74103 respectively. The Language variable had the highest loading on Function 2. The Concepts and Motor variables had the greatest influence on Group 1 membership which was negatively influenced by Language. Group 3 membership was most influenced by the Concepts and Motor variables and negatively influenced by Language.

In GSRT analysis, the Maturity variable loaded highest on Function 1 with a loading of 0.99947. The Achievement variable had relatively high loadings on Function 1 (0.78777) and Function 2 (0.61597), while Maturity had a low discriminant loading (0.03246) on Function 2.

The influence of GSRT predictor variables at the third grade varied across groups. Group 1 members had low scores on both Maturity and Achievement while Group 2 subjects scored higher on Achievement and lower on Maturity. In contrast, Group 3 subjects scored higher on GSRT Maturity and lower on GSRT Achievement.

At the fifth grade, GSRT Maturity was more evenly distributed between Function 1 (0.72521) and Function 2 (0.68853). Achievement had a discriminant loading of 0.98911 on Function 2. Group 1 subjects scored higher on Function 2 (a combination of Maturity and Achievement) while Group 3 subjects scored highest on Function 1 with a high loading on Maturity.

In the third grade PSB analysis, PSB Language and Cognition loaded heavily on Function 1. Cognition and Perception were more evenly distributed on Function 2 while PSB Language loaded negatively on Function 2.

Group 1 subjects scored lowest on PSB Language, and performed higher on Cognition and Perception. Group 2 membership appeared to be influenced more by Cognition and Perception, and Group 3 membership was influenced most by PSB Language.

At Grade 5, this pattern was repeated with a high PSB Language loading (0.94127) on Function 1, and a moderate discriminant loading (0.68759) for Cognition. Perception was evenly distributed between Functions 1 and 2. Colors and Perception loaded on Function 2 (0.70850 and 0.53357 respectively).

Again, Group 1 subjects scored lowest on PSB Language and Cognition and higher on Perception and Colors. Group 2 subjects scored higher on Perception and Colors and lower on Language and Cognition while Group 3 subjects scored highest on Function 1 which loaded heavily on Language and Perception.

Overall analysis of the three batteries resulted in the identification of three constructs, a nonverbal construct related to physical maturity (Motor-DIAL-R, Perception-PSB, and Maturity-GSRT), a Conceptual construct related to cognitive skills (Concepts-DIAL-R and Cognition-PSB), and a Language construct (Language-DIAL-R, Language-PSB and Achievement-GSRT). Higher nonverbal scores and lower language and conceptual scores were related to Group 1 membership. Group 3 membership was related to higher language and conceptual scores and lower non-verbal scores. Skills tapped on the Colors subtest influenced Group 2 membership most.

Demographic Differences

In the DIAL-R analysis, females performed slightly better than males on Function 1 which had moderate to high loadings on all three subtests, and white subjects performed better than non-white subjects. Independent t-tests for Independent Samples (using the Compare Means Option, Statistics procedure of SPSS for Windows Advanced Version V6.1 (SPSS, 1995) did not yield significant gender differences, however.

The GSRT analysis resulted in gender differences on Function 1 with males scoring slightly higher. This finding does not corroborate the popular belief that females mature more rapidly than males. Males scored higher than females on all GSRT subtests including the composite variable, GSRT Maturity, which had high loadings on Function 1. Significant differences were found on Maturity (t-value

1.99, p.= 0.05), Interview (t-value 3.14, p.= 0.002), and Naming (t-value 2.02, p.= 0.047).

On the PSB analysis, females scored higher than males on Function 1 which had high loadings on Readiness, Blocks, and General Information. Higher SES scores (0.74750) were also associated with higher scores on Function 1. Function 2, with high loadings on Visual Discrimination, Auditory Memory, and Visual Memory was negatively influenced by SES (-0.13398) and ethnicity (-0.14237). Function 3 with high loadings on Colors (0.71230), Commissions (0.47036), and Auditory Memory (0.34698) was influenced by gender (0.7521) with females performing better than males. These gender differences were substantiated by Independent t-tests which resulted in significant gender differences for Colors (p = 0.01), Commissions (p = 0.06), and Auditory Memory (p = 0.05).

Posteria Probabilities

Posteria probability estimates for individual subjects were examined for outliers and in-doubt subjects. These estimates were provided by the classification function of SPSS for Windows Advanced Version V6.1 (SPSS, 1995) and reflect the "typicalness" of a subject's scores on the predictor variables (or score vector) to scores of other subjects in the same group (Huberty, 1994, p. 76). The score is actually the proportion of subjects in the group who have scores (score vectors) close to the subject in question (Huberty, 1994, p. 76). Subjects with typicality probabilities less than 0.10 were examined to detect patterns of predictor variable scores or discriminant function scores which would set the subject at question apart from other subjects in the group (Huberty, 1994, p. 94). One-Sample t-tests were computed by using the Compare Means option of the Statistics function SPSS for Windows Advanced version V6.1 (SPSS, 1995) to test for significant differences between the mean scores on predictor variables for the nontypical subjects and the mean scores of the total group (Norusis, 1993). Scatterplots were also examined for outliers.

Eleven DIAL-R subjects had typicality probability scores below 0.10. Of the 11 subjects, two were in Group 2 and nine were in Group 1. One Sample t-tests were computed comparing nontypical Group 1 means for predictor variables with total Group 1 means. Significant differences were obtained on all three subtests with the nontypical group scoring lower than the group centroids. (See results in Table D-1, Appendix D). The typical pattern for this group was to score highest on the Language subtest, followed by the Motor subtest, and lowest on Concepts.

The nontypical Group 2 subjects were similar to the nontypical Group 1 subjects on the Motor variable (17.33 and 17.00) and on Language, 22.11 and 23.00). However, the nontypical Group 2 Concepts mean (25) was considerably higher than the nontypical Group 1 Concepts mean (12.00). This difference in Concepts variable means appears to be the discriminating factor between these two nontypical groups.

Nine Gesell subjects had typicality probability scores below 0.10 (3 in Group 1, 2 in Group 2, and 4 in Group 3). Non-typical Group 1 subjects differed on Writing Name and Numbers, and Naming Animals and Interests. Non-typical Group 2 subjects differed on Cubes, Incomplete Man, Naming Animals and Interests, and the Interview subtests. Group 3 differed on Cubes, Copyform, Writing Name and Numbers, and the Interview subtest. (See results in Table D-2 in Appendix D).

Four PSB subjects had typicality scores below 0.10 (1 in Group 1, 2 in Group 2, and 1 in Group 3). The three subjects in Group 1 and Group 2 had non-typical responses to at least one of the subtests (raw scores of 0). The Group 3 subject's receptive and expressive language scores were unusually high when compared with other subtest scores.

Although significant differences were obtained between the non-typical subjects and typical subjects on some predictor variables, a clear pattern which would substantiate the existence of a fourth subgroup was not identified. As a result, the nontypical subjects were included in all analyses.

Classification Errors

Two additional subgroups were identified when the individual unit results were reviewed: (1) subjects with predicted Group 1 membership who had actual Group 3 membership, and (2) subjects with predicted Group 3 membership who had actual Group 1 membership. Subgroup totals are listed in Table 18.

Table 18

Classification Errors Subgroup Totals

	Group 1 to 3	Group 3 to 1
DIAL-R	4	5
GSRT	9	1
PSB	2	1

1 to 3 Group. The 1 to 3 group was comprised of four DIAL-R subjects, nine GSRT subjects and two PSB subjects. Mean subtest scores for DIAL-R subjects did not differ significantly from other subjects in Group 3 on the Concepts and Language subtests. The mean score for the 1 to 3 group

was higher on the Motor subtest than others in Group 3, however. T-test values are listed in Table E in Appendix E.

Individual differences related to medical problems and handicapping conditions were noted. One subject had asthma and another had been placed in a developmental preschool before being screened. Another subject was six years old at the time of screening. Group 1 to 3 DIAL-R classification errors appeared to have been the result of individual subject characteristics.

Two notable differences were found in the Gesell sample between subjects in the 1 to 3 group and the total Gesell group. First, the retention rate for Group 1 to 3 (66%) was noticeably larger than the total Gesell Group retention rate (32%). Secondly, when the subtest means for Group 1 to 3 subjects were compared with means from the total Gesell group, the Group 1 to 3 means were significantly lower on the Cubes subtest (at the third grade), and on the Naming Animals and Interests subtest at third and fifth grades. See t-test values in Appendix E.

Two PSB subjects in the 1 to 3 group were significantly higher on expressive language, as measured by the EOWPVT, than other PSB subjects. These higher scores seemed to have placed the PSB subjects in a higher predicted reading group.

<u>3 to 1 Group.</u> Differences among the 3 to 1 group will be discussed next. The PSB subject had a low total PSB

score (24) compared to the PSB mean total score (48.58). This student was recommended for referral to an early intervention program at the local mental health center as a result of the screening session. Comments indicated that the subject did not respond to all test items and that the results should have been considered questionable. The descriptive data on this subject suggested that this incorrect classification resulted from invalid preschool data.

The GSRT subject, an academically gifted student, appeared to have the appropriate actual group placement. This subject's score on the Cubes subtest (4 years 6 months) was noticeably lower than his mean score for all subtests (5 year 5 months). This relatively lower Cubes score appears to have influenced predicted group membership negatively.

An uneven pattern was also observed among the DIAL-R subjects in the 3 to 1 group. Three out of the five subjects had one subtest score which was considerably lower than others. The other two subjects had low total scores.

<u>Regression Analysis</u>

Supplementary regression analyses were computed to estimate the relationship between predictor variables and the reading subskills measured by the California Achievement Test. The relationship between predictor variables from the

DIAL-R, GSRT, PSB, and the Vocabulary, Comprehension, and Word Analysis subtests of the California Achievement Test were explored through a series of multiple regression analyses. The procedures used for entering variables were: direct, backward, and forward methods. The criterion used for entering a variable was the probability of F-to-enter with a default value of 0.05. In this case, a variable was entered when it had a probability associated with the F test less than or equal to 0.05 (Norusis, 1993). When the Backward method was used, the default criterion of a maximum probability of 0.10 was used to determine eligibility for removal. Table 19 includes standardized beta weights for predictor variables from the regression analyses, and the adjusted R squared coefficients for the best set of variables from each analysis. According to Norusis (1993), beta coefficients may be interpreted in "approximate percentage terms" (p. 342).

The results of the analyses using different methods (direct, backward, and forward) were quite similar. Because of this, the beta weights for the direct method (considered representative of the analyses) were selected for reporting.

Partial correlations were also computed between predictor variables and each specific reading subskill (controlling for gender, SES, ethnicity. The purpose of these analyses was to determine the significance of

Table 19

Beta Coefficients and Adjusted R Squared Coefficients From Regression Analysis Between Predictors and Reading Subskills

DIAL-R		Gesell		PSB	
Word Analysis					
Concepts	0.367	**Copy Forms**	0.312	**Cognition**	0.217
Motor	0.189 (0.35)	**Cubes**	0.219 (0.19)	**EOWPVT**	0.205
Language	0.100	Inc. Man	0.062	**Colors**	0.144 (0.25)
		Interview	0.011	Perception	0.126
		Naming	0.042	PPVT-R	0.027
		Nam/Num	0.149		
Vocabulary					
Concepts	0.484 (0.37)	**Copy Forms**	0.282	**EOWPVT**	0.286
Motor	0.099	**Cubes**	0.271	**PPVT-R**	0.194
Language	0.094	**Inc. Man**	0.249 (0.19)	**Colors**	0.162 (0.25)
		Nam/Num	0.060	Cognitive	0.141
		Interview	0.113	Perception	0.067
Comprehension					
Concepts	0.451 (0.29)	**Inc. Man**	0.460 (0.15)	**Colors**	0.256
Language	0.101	Cubes	0.039	**EOWPVT**	0.240
Motor	0.040	Interview	0.021	**Perception**	0.168 (0.35)
		Copy Forms	0.017	Cognition	0.133
		Nam/Num	0.016	PPVT-R	0.099
		Naming	0.179		

Note: variables in the best set of predictors are in bold print, and Adjusted R squared coefficients for best set of variables are listed in parenthesis.

predictor variables when the variations attributed to gender, ethnicity, and SES were partialled out.

Multiple Regression Analysis. Concepts accounted for the largest proportion of variance for all three outcome measures in the DIAL-R analysis. (See Table 19 for beta weights and adjusted R squared coefficients). The Motor subtest (0.189) influenced Word Analysis more then Comprehension (0.040) or Vocabulary (0.099). Vacc et al. (1987) also found the DIAL Concepts and Fine Motor subtests to be effective predictors of reading performance at the first grade.

In order of importance, the Gesell Copy Forms, Cubes and Incomplete Man subtests accounted for the largest proportion of variance in Vocabulary and Word Analysis. The pattern changed slightly on the Comprehension analysis with Incomplete Man accounting for the most variance (0.46).

The EOWPVT accounted for the largest proportion of variance among the PSB predictor variables on Word Analysis, (0.205), Vocabulary, (0.286), and Comprehension, (0.256). Vocabulary was the only outcome variable influenced significantly by the other language measure, PPVT-R.

Cognition accounted for more of the variation in Word Analysis than any other PSB predictor (0.217) yet had little influence on the other two outcome measurements.

The most noticeable finding from Table 19 is the consistent pattern of influence attributable to a few predictor variables across preschool screening batteries. In DIAL-R and GSRT analyses, the most important predictors were Concepts, Copy Forms, Cubes, and Incomplete Man. In the PSB analyses, EOWPVT and Colors were included in the best set of predictors for all three outcome measures.

Despite this commonality, some variations occurred among predictor variables when they were regressed on the reading subskills. As expected, Word Analysis was influenced by Concepts, expressive language (EOWPVT), and predictors which purport to measure physical or behavioral maturity, the DIAL-R Motor, GSRT Copy Forms, and the GSRT Cubes subtests.

The DTVMI was entered into the regression analysis with other PSB variables to further test the hypothesis that Word Analysis is influenced by a physical maturity factor. In this analysis, the DTVMI was included in the best set of predictors for Word Analysis, but was not an influential predictor of other outcome measures.

Vocabulary was influenced similarly by Concepts, Copy Forms, Cubes, Colors, and expressive language (EOWPVT). Cognition dropped out of the best set of predictors, however, being replaced by the PPVT-R, and the Incomplete

Man variable which was added to the Gesell best set of predictors.

Comprehension was influenced by the Concepts, Incomplete Man, Colors, EOWPVT, and Perception predictors. The PPVT-R was replaced by PSB Perception and Cubes and Copy Forms dropped out of the best set of Gesell predictors when associated with Comprehension.

Partial correlations. When predictors were correlated with each specific reading scale, with the effects of demographics partialled out, most of the predictor variables made significant contributions at the 0.05 level to each reading scale. The exceptions were the Naming variable in the Gesell analysis which did not make significant contributions to any of the reading subskills and the Writing Name and Numbers variable which correlated significantly with Word Analysis at the 0.025 level and Vocabulary at the 0.009 level, but was not a significant contributor to Comprehension (0.089).

In the Preschool Screening Battery analysis, the Colors variable did not contribute significantly to Word Analysis or Vocabulary, but was significantly related to Comprehension (0.1977. p. = 0.05). This finding suggested that the Colors variable accounted for a unique proportion of the variation in Comprehension which was independent of the influence of demographic variables.

A series of partial correlations were computed in an effort to determine what, if any, influence the Colors subtest had on total reading achievement at third and fifth grades with the effects of ethnicity, gender, and SES partialled out. In each case, the Colors variable was significantly correlated with the Reading 3 variable and Reading 5 variable. When controlling for all three demographic variables, the correlations between total reading achievement and Colors were 0.1897 (p. = 0.060) at third grade and 0.1728 (p. = 0.087) at the fifth. Similar results were obtained when gender and SES were controlled (0.1835, p. = 0.068 and 0.1692, p. = 0.092).

When controlling for SES alone, the correlations increased to 0.2561 and 0.2337 (significant at the 0.01 level at third grade and at 0.02 level at the fifth grade). When the effect of gender was partialled out, the correlations were 0.2666 and 0.2795 which were significant at the 0.01 level. In the final analysis when the effect of Ethnicity was partialled out, the correlations between reading achievement and Colors were 0.3216 at third grade and 0.3209 at fifth grade (both significant at the 0.01 level).

Additional partial correlational analyses were computed controlling for demographic variables and the other two specific reading subskills. The purpose of these analyses was to determine which, if any, of the predictor variables made unique contributions to a specific reading subskill when the effects of demographic variables and the other two subskills were partialled out. This step was included because of high intercorrelations among the subtests and the total reading score of the California Achievement Test which ranged from 0.7150 to 0.9481.

The results of these analyses indicated that Word Analysis was influenced by the Gesell Cubes variable which accounted for 5.9% of the variance (p. = 0.006) with the effects of demographics, Vocabulary, and Comprehension removed. The Writing Name and Numbers variable accounted for 1.18% (p. = 0.0668) of the variance. The DIAL-R Motor variable accounted for 2.23% (p. = 0.091) of the variance in Word Analysis, and the DIAL-R Language variable and the DTVMI accounted for 2.09% (p. = 0.166), and 1.7% (p. = 0.194 respectively). This pattern of influence was similar to that found in the multiple regression analyses reported in Table 19.

Vocabulary was influenced by the DIAL-R Concepts, PSB EOWPVT, and PSB PPVT-R variables. These predictors accounted for 4.43% (p. = 0.043), 6.6% (p. = 0.014), and 8.68% (P. = 0.003) of the variation in Vocabulary when demographics, Word Analysis, and Comprehension were partialled out.

Comprehension was influenced significantly by the Cubes variable when demographics, Word Analysis, and Vocabulary were partialled out (-0.2644, p. = 0.032). Note that the correlation between Cubes and Comprehension was negative.

The PSB Colors variable accounted for 1.78% (p. = 0.192) of the variation in Comprehension and the PSB Perception variable accounted for 2.1% (p. = 0.159). Additional partial correlations were computed between Auditory Memory, Visual Discrimination, Visual Memory, and Commissions (PSB predictors included in the PSB Perception Variable), and Comprehension.

As a result, Auditory Memory was the only Perception predictor which contributed significantly to Comprehension when the effects of the other three Perception predictors were partialled out (r = 0.2358, p. = 0.020). Auditory Memory continued to contribute significantly when the effects of other Perception variables, demographics, and other subskills were partialled out (r. = 0.2155, p. = 0.037). Stevenson and Newman (1986) obtained a correlation of 0.43 between a similar task (serial memory for numbers) and reading comprehension at the tenth grade.

In summary, word analysis skills were influenced by conceptual development (knowledge and understanding of basic concepts), a mixture of fine motor skills including eye-hand coordination, visual perception, and expressive language skills. Vocabulary skills were influenced by conceptual development, expressive and receptive language skills, knowledge of colors, and abstract thinking skills represented by the Incomplete Man tasks. (The interpretation of the Incomplete Man subtest as an abstract thinking task was presented by Kaufman (1971) who hypothesized that the Incomplete Man tasks required understanding of part-whole relationships). It is presumed that providing the missing part or utilizing knowledge of part to whole relationships is a higher level thinking skill than reproduction of a geometric symbol.

In the PSB analysis, Comprehension was influenced by the Colors (0.1336, p. = 0.192) and Perception (0.1441, p. = 0.159) variables when demographics, Word Analysis, and Vocabulary were partialled out. Auditory Memory emerged as the single significant contributor to Comprehension when the effects of the other Perception predictors were partialled out (0.2358, p. = 0.020).

The probability values associated with correlations between the remaining predictor variables and Comprehension (with demographic variables, Word Analysis, and Vocabulary partialled out) ranged from 0.304 to 0.975.

Overall, when the effects of demographic variables and other reading subskills were partialled out, the resulting correlations between predictor variables and reading comprehension were small. These small correlations suggested that the reading subskills measured are intercorrelated, and may serve as intervening variables between the predictor variables and higher level reading skills.

The results of the zero order correlations and partial correlations converged resulting in configurations of influence similar to the developmental trend in reading acquisition described by Stanovich (1986). Stanovich (1986) observed that the skills and abilities presumed to be precursors of reading achievement became "increasingly interrelated as the age of the children increased" (p. 296).

Summary

The results of the study were presented in this chapter. A 54% return rate was obtained from letters requesting parental consent. Of those who responded positively, 296 were used in discriminant function analysis to predict group membership at the third and fifth grades. Predictive discriminant analysis resulted in better than chance classification accuracies with the DIAL-R and the PSB at Grades 3 and 5. The GSRT analysis results were not better than chance except at predicting Group 2 membership when demographic variables were included at the fifth grade. The PSB analysis resulted in medium effect sizes (0.47 and

0.53) at the third and fifth grades. Other effect sizes were small. Three factors extracted from factor analysis and predictive discriminant analysis which influenced reading achievement significantly across batteries were: (1) a non-verbal construct, (2) a conceptual construct, and (3) a language construct. A few predictor variables accounted for the majority of the variation in reading subskills measured on the California Achievement Test. Notable differences were found in the retention rates of Gesell subjects and predictor profiles of DIAL-R and Preschool Screening Battery subjects in the 1 to 3 and 3 to 1 groups when compared with total group results.

CHAPTER V

Discussion and Recommendations

The results of the study are summarized and discussed in this chapter. First the methodology employed is reviewed. Next, the summarized results and conclusions related to the research questions are presented. Finally, implications of the findings as they relate to the four research questions, and recommendations for future research are discussed.

Review of Methodology

The purpose of this study was to assess the predictive validity of three preschool screening batteries, and to identify any early determinants which may have influenced the acquisition of reading achievement from school entrance to fifth grade. Subjects were sixth grade students from five school systems in North Carolina. Parental consent forms were sent to parents of all sixth grade students in the participating school systems. A return rate of 54.7% was obtained. Of the 575 returned consent forms, 466 (44%) gave permission for the record review. That number dropped to 296 subjects because 170 of the subjects with yes responses did not have preschool screening data in their school folders. The sample size for each preschool screening battery was Dial-R (n=101), PSB (n=102), and GSRT (n=93). Data were collected through record reviews at the individual schools, and follow-up visits and telephone calls were used to obtain missing data.

Statistical analyses were conducted using SPSS for Windows Advanced Version V6.1 (SPSS, 1995). Principal component analysis and principal axis factoring procedures were conducted to reduce the number of variables for the Preschool Screening Battery and Gesell analyses. Reliability coefficients were computed to assess the reliability of the combined predictor variables. Six predictive discriminant analyses were computed with predictor variables and grouping variables for third and fifth grades. Next, three demographic variables were included in the predictive discriminant analyses to control for gender, SES, and ethnicity. Demographic trends were examined in further exploratory discriminant function and correlational analyses. Independent t-tests were computed to identify significant gender differences.

Statistical tests were calculated to determine if classification accuracies obtained from predictive discriminant analyses were better than chance. Press's Q statistics and z-scores were calculated (a procedure in which hit rates from the analyses were compared with the proportional chance criterion and maximum chance criterion to determine significance). Mean scores were compared by t-tests to determine if significant differences existed between predictor variables of correctly classified subjects and those with classification errors.

Regression analyses were computed to assess the relationship between predictor variables and reading subskills measured on the California Achievement Test (Word Analysis, Vocabulary, and Comprehension). Direct, forward, and backward stepping methods were used to determine the best set of predictors for each battery and each outcome measure. Partial correlations were computed to determine the power of predictor variables to influence specific reading subskills when the effects of demographic variables and other reading subskills were partialled out.

Summary of Results and Conclusions

The results of the study will be summarized and discussed as they relate to each of the research questions.

1. What is the predictive validity of the Gesell School Readiness Test (GSRT), the Developmental Indicators for the Assessment of Learning (DIAL-R), and the Preschool Screening Battery (PSB).

Predictive discriminant analyses (PDA) computed with Preschool Screening Battery variables resulted in better than chance classification accuracy on 11 out of 12 statistical tests computed. The probability value of the

12th test (which was for Group 2 in the PSB fifth grade analysis) was very close to reaching significance (0.053). The DIAL-R analyses resulted in significance on 10 out of 12 tests, and the Gesell attained significance on 3 tests out of 12.

The Preschool Screening Battery attained the highest overall hit rates, 62.75% at the third grade and 70.00% at the fifth grade compared with 55.00% and 59.18% for the DIAL-R and 43.48% and 44.44% for the Gesell. The DIAL-R had the highest hit rates at third grade for Group 1 membership, and the Preschool Screening Battery had the highest hit rates for Group 1 at the fifth grade and for Group 3 at the third and fifth grades.

Separate group hit rates for Group 1 were of particular interest as early intervention strategies should be directed toward subjects who may be at-risk for below average reading achievement. The separate group hit rate computed for the DIAL-R Group 1 at the third grade was 90%. The Preschool Screening Battery and the Gesell hit rates for Group 1 (37.59 and 28.60) were lower than the DIAL-R hit rate. The Preschool Screening Battery Group 1 hit rate was significant at the 0.05 level, however.

The Preschool Screening Battery had the highest fifth grade Group 1 hit rate (66.70%) compared to the DIAL-R rate (56.60%). The Gesell had the lowest hit rates for Group 1

(25%). For clarity, probability values for z-scores and Press's Q statistics are included in Table 21.

Although no guidelines are available to determine how high classification accuracy should be relative to chance, Hair et al. (1992) suggested using a 25% improvement over the maximum chance criterion as a cutting line to determine adequate classification accuracy. When compared with Hair's criterion, all three batteries were 25% better than the maximum chance criterion. It is noteworthy that the Gesell analyses were only marginally better than chance. See Classification ratios compared with Hair's criterion in Table 20.

Table 20

Classification Accuracy Percentages Compared with 25% Improvement Over Maximum Chance Criterions

Group	25% Improvement	Actual Classification Accuracy	
		Third	Fifth
DIAL-R	51%	55.00%*	59.19%*
GSRT	44%	43.48%	44.44%*
PSB	58%	62.75%*	70.00%*

Note: *= Percentages above maximum chance criterion plus 25%.

Table 21

Summary of Statistical Tests of PDA at 3rd and 5th Grades

Group	Grade	DIAL-R	Gesell	PSB
Z Score P-value				
1	3rd	0.018	NS	0.044
	5th	0.000	NS	0.000
2	3rd	0.034	NS	0.030
	5th	0.043	0.000	0.053
3	3rd	NS	0.014	0.000
	5th	0.000	NS	0.000
Total	3rd	0.000	NS	0.000
	5th	0.000	NS	0.000
Press's Q Statistics				
Grade		DIAL-R	Gesell	PSB
3rd	Without Demographic	10.44	NS	19.85
	With Demographic	17.15	NS	19.85
5th	Without Demographic	14.73	NS	24.51
	With Demographic	NS	6.40*	15.21

When the reduction in error rates (I-values) were computed, the Preschool Screening Battery had medium effect sizes at third grade and at fifth grade. These effect sizes (0.47 at 3rd and 0.53 at 5th) may be interpreted as a 47% to 53% reduction in classification errors when the Preschool Screening Battery was used to classify subjects. In comparison, the Gesell and DIAL-R had small effect sizes. Table 16 contains of I-values.

2. What are the early determinants identifiable through discriminant function analysis of these test batteries which contribute significantly to reading achievement across grade levels.

Three early determinants of reading achievement were identified through predicative discriminant analysis and factor analysis: a non-verbal determinant, similar to Gesell's construct, behavioral maturity (Kaufman, 1971), a conceptual determinant, and a language comprehension determinant.

These results were similar to those reported in previous research. Vacc et al. (1987) found the Concepts subtest to be the best DIAL-R predictor of reading achievement at the first grade. The language construct is similar to that described by Stanovich et al. (1984) which had high loadings on the PPVT and a listening comprehension task.

In the present study, a pattern emerged across preschool screening batteries and across grade levels in which the construct influencing predicted group membership most at the third grade also exhibited the most discriminating power at the fifth grade for above average readers. These constructs were: DIAL-R Concepts, Gesell

Maturity, and the Preschool Screening Battery Language variable. These findings were similar to the trend reported by Stanovich et al. (1984) in which the early skills and abilities measured became "increasingly interrelated as the age of the children" increased (p. 296).

The discriminatory power of the constructs varied among the three reading achievement groups, and the following discussion will be based on two assumptions. First, reading acquisition is a process (Snider & Tarver, 1987), and second, the three reading achievement groups in this study represent a continuum of reading skills which change as reading competencies become more advanced (Fletcher & Satz, 1980; Stanovich et al., 1984).

Based on these assumptions, the effects of the constructs across grade levels and reading groups will be reported. In the DIAL-R analysis, motor skills influenced beginning readers, but had diminished influence on average readers while conceptual and language skills increased in importance with the average group. This trend continued until the influence of motor skills was negligible. Sears and Keogh (1993) reported similar trends where early measures of visual processing were robust at the early stages of reading acquisition, but were less important at later stages.

At the fifth grade level, the Language subtest gained in influence with Groups 1 and 2, and Concepts had the greatest influence on Group 3 readers. Building on this, one might hypothesize that the constructs included in the DIAL-R analysis follow this pattern of influence: physical maturity being most important at the beginning stages of reading acquisition, followed by language skills in the intermediate stages, followed by concepts and cognitive skills at the comprehension level.

The Language variable was least effective in predicting reading achievement of the three DIAL-R predictors. These findings supported the work of Vacc et al. (1987) who studied the DIAL (predecessor to the DIAL-R) and found that the best model for predicting reading achievement at the first grade included the Concepts and Fine Motor subtests and excluded the Language subtest. In the PSB third grade analysis, PSB Cognition and PSB Perception were associated with beginning reading skills (Groups 1 and 2) followed by the emergence of Language as the most important influence with Group 3. At the fifth grade, PSB Perception and Colors were associated with below average and intermediate reading skills with Language and Cognition emerging as the primary influence on reading achievement with above average readers. Again the same path of influence emerged: biological maturity followed by language development followed by a combination of language and cognition. This pattern of influence is suggestive of the multiple cause theory of reading discussed by Stanovich et al. (1984).

3. Does the effect of the determinants vary or change when they are associated with different outcome measurements such as word analysis, vocabulary, or reading comprehension?

Supplementary multiple regression analyses were computed to assess the relationship between preschool predictors and the following California Achievement Test subtests: Vocabulary, Word Analysis, and Comprehension. Predictors from each preschool screening battery were regressed on Vocabulary, Word Analysis, and Comprehension in three separate analyses. As a result, eight significant predictors emerged from a total of 14 predictor variables. The predictors are listed in Table 22.

Table 22

Significant Predictor Variables From Multiple Regression Analysis

Group	Variables				
DIAL-R	Concepts				
GSRT	Inc. Man	Copy Forms	Cubes		
PSB	EOWPVT	Colors	Cognition	Perception	PPVT

Unique patterns were identified for different outcome measures. Word Analysis was influenced by Concepts, Copy Forms, Cubes, Cognition, EOWPVT, and Colors. These results suggested that all three early determinants accounted for a significant amount of variance in Word Analysis (conceptual/ language development, fine motor skills, and visual perception).

Vocabulary scores were influenced by Concepts, Copy Forms, Cubes, Incomplete Man, EOWPVT, PPVT-R, and Colors. The inclusion of the PPVT-R and Incomplete Man variables suggested, as expected, that vocabulary skills measured on the Vocabulary subtest are heavily dependent on receptive and expressive language skills and are more abstract than word analysis skills.

The Comprehension subtest was influenced by Concepts, Incomplete Man, Colors, EOWPVT, and Perception. Fine motor skills measured on the Gesell Battery dropped out of the best set of predictors on the Comprehension analysis indicating these skills were not pertinent to comprehension tasks while the PSB Perception variable was. (The Perception variable is a combination of auditory memory, visual discrimination, and listening tasks.) Comprehension was the only outcome measure influenced significantly by Perception. Of the four predictors which were combined into the Perception variable, Auditory Memory was the only one to emerge as a significant predictor of Comprehension skills. These results suggested that Comprehension was affected by language skills, abstract and associative thinking skills, and auditory memory skills.

Additional partial correlations were computed with the effects of demographic variables and other reading subskills partialled out. The influence of most predictors was enhanced when the effects of demographics alone were partialled out. The exceptions were the Gesell variable, Naming, which was not a significant contributor to any of the subskills, and the Writing Name and Numbers variable which did not contribute significantly to Comprehension. These two variables were not significant contributors in the zero order correlations either.

When the effects of the other two reading subskills and demographics were partialled out, the unique Word Analysis variance associated with the Gesell Cubes variable was significant. The unique Vocabulary variance associated with DIAL-R Concepts, Gesell Cubes, PSB EOWPVT and the PSB PPVT-R was also significant. The PSB Colors and the Perception variables accounted for 1.78 and 2.1% of the unique variance associated with Comprehension even though they were not significant at the 0.05 probability level.

In efforts to explain the relationship of the Colors variable to reading achievement in the fifth grade predictive discriminant analysis, partial correlations were computed between Colors and Reading 3 and Reading 5. Three analyses were computed controlling for gender, SES, and ethnicity. When the effect of SES was partialled out, the Colors variable was associated with Reading 3 and Reading 5 at the 0.019 and 0.000 levels. When the effect of the Gender variable was partialled out, the Colors variable was associated with Reading 3 and Reading 5 at p-values of 0.007 and 0.005. The effect of ethnicity was partialled out with similar results (0.001 and 0.001). When the effects of SES, gender, and ethnicity were partialled out simultaneously, the unique variance in Reading 3 and Reading 5 associated with Colors dropped to 0.1897, p. = 0.06 and 0.1728, p. = 0.087.

In conclusion, the Colors variable seems to be associated with unique reading variance which is independent of SES, gender, and ethnicity. These results do not explain why the Colors variable was important in improving classification results for Group 1 at the fifth grade, but did not improve classification accuracy at the third grade.

4. Are there significant differences in groups of children who were correctly identified as not at-risk by preschool screening data, and those who were classified as not at-risk, but who developed reading disabilities later?

Incorrectly classified subjects were designated as those with at least one group level between actual group placement and predicted group membership (1 to 3 or 3 to 1).

Subjects in the 1 to 3 group had predicted Group 3 membership (based on preschool screening data) yet had actual Group 1 membership. The Gesell analysis had seven subjects in the 1 to 3 group at the third grade. Significant differences were found in the Naming Animals and Interests and the Cubes subtests (p-value = 0.01) between subjects in the 1 to 3 group and correctly classified Gesell subjects (the Cubes and Naming Animals and Interests subtests were lower).

Four of the Gesell Group 1 to 3 subjects had been enrolled in a transitional kindergarten or transitional kindergarten-first grade class. Despite being given the extra year for catch up (a practice described by May & Welch, 1984), these four subjects had reading skills at fifth grade which placed them in the below average reading group. These findings supported those of May and Welch (1984) who found that children judged to be immature on the Gesell screening, and who consequently spent an extra year at home before starting school or spent an extra year between kindergarten and first grade, scored lower than other children on the Stanford Achievement Test and the New York State Pupil Evaluation Program test at 4th and 6th grades.

In contrast, these results do not support research conducted by Ames and Ilg (1964) and Wood et al. (1984) who claimed performance on the Gesell was a good predictor of school performance. In the Ames and Ilg (1964) study, school achievement was determined by general school performance, but the method for determining the level of school success was not described. In the Wood et al. (1984) study, readiness to begin school was determined by developmental age on the Gesell, and failure in kindergarten was determined by special needs status. In both of these studies, academic success was determined by methods different from those used in this study.

Four DIAL-R subjects in the 1 to 3 group had significantly higher scores on the Motor subtest than other Group 3 subjects. Significant differences were not found in the Concepts and Language scores.

Relatively higher scores on expressive language measures appeared to inflate the predicted reading group membership of the two Preschool Screening Battery subjects in the 1 to 3 group. These subjects had language ages of 6 years 8 months and 7 years at the time their chronological ages were 5-0 and 5-1.

Characteristics which appeared to influence (3 to 1) group predictions in the Preschool Screening Battery and the DIAL-R analyses were: chronic illness, handicapping conditions, and behavioral problems. Intra-subtest scatter noted on the DIAL-R profiles appeared to influence inaccurate classification. The subjects each had one predictor variable which was 3 to 8 points lower than the other two predictors. The one Gesell subject who placed in the 3 to 1 group was later identified as Academically Gifted. This student had a relatively low score on the Cubes subtest which was the most significant Gesell predictor.

Discussion

The purpose of this study was to assess the predictive validity of three preschool screening batteries, the DIAL-R, the Gesell, and the Preschool Screening Battery, and to identify any underlying constructs of the preschool batteries which acted as early determinants of reading acquisition. Although sample sizes for the low reading achievement groups were small for the Preschool Screening Battery and the Gesell, the n:p ratios were large enough for the analyses to be considered valid. Overall group sizes and demographics for the total sample were representative of sixth grade students in North Carolina. Total group sizes for each battery were considered adequate for discriminant analyses (DIAL-R, n= 102, GSRT, n= 93, PSB, n= 102). The research questions of this study will be used as a framework for the remainder of this discussion.

Predictive Validity

When classification accuracy ratios were compared with a maximum chance criterion plus 25% improvement criterion, the discriminatory power of all three batteries was considered better than chance. However, when a more stringent method was used (standard normal statistics) a procedure in which the number of correctly classified cases was compared with the proportional chance criterion, the results were different. Preschool Screening Battery and DIAL-R analyses yielded significant overall results at the third and fifth grades. The Gesell results were not significantly better than chance except with a very specific situation (average students at the third grade when demographics were included).

The discriminatory power of the preschool batteries to classify subjects in the low reading achievement group (Group 1) was of particular interest to this study. The DIAL-R and the Preschool Screening Battery both proved to be better than chance at classifying Group 1 subjects. This suggested that the DIAL-R does what it purports to do, identify children who have potential problems (Johnson & Goldman, 1990). The DIAL-R's power to classify above average readers was not as accurate (better than chance at fifth but not the third grade).

The Preschool Screening Battery analyses were significantly better than chance at classifying subjects in all groups. Classification ratios at third and fifth grades were 34% and 48% better than the maximum chance criterion. The Preschool Screening Battery effect sizes (0.47 and 0.53) were also superior to those obtained from DIAL-R and Gesell analyses. The Preschool Screening Battery also had classification accuracies of 75% and 81% with Group 3 subjects.

Based on the findings of this study, the discriminatory power of the DIAL-R and the Preschool Screening Battery should be considered sufficient to warrant their use in the early identification of subjects with potential reading problems.

The discriminatory power of the Gesell, however, does not support its use as a predictor of future reading problems. This data raises doubts regarding the validity of the Gesell as a measure of readiness for school entrance, and the premise that school placement should be based on developmental age rather than chronological age (Ames & Ilg, 1964; Wood et al., 1984).

Based on the supposition that learning to read is the most important task a student faces during the first three years of elementary school, how can the Gesell be considered a useful predictor of future achievement if it does not measure reading readiness to a significant degree?

Early Determinants of Reading Acquisition

Predictor variables formed from combinations of subtests of the preschool batteries were roughly equivalent to the following categories: non-verbal (neurological/ perceptual), conceptual, and language based. These categories were similar to those used in previous research (Tramontana et al. 1988).

The Concepts subtest was the most important variable in the DIAL-R battery, and ranked in importance with the Preschool Screening Battery Language variable. The Gesell variable, Maturity, contributed significantly to Word Analysis, and when the effects of demographics, Vocabulary, and Comprehension were partialled out, the unique Word Analysis variance associated with the Gesell Cubes subtest continued to be significant.

Surprisingly, the PSB Colors variable emerged as an important contributor to prediction of reading achievement at the fifth grade level. Correlations between the Colors variable and Gender, Ethnicity, and SES were computed to determine if the Colors variable was related to demographic variables specifically, SES. Significant correlations (0.29, $p. = 0.003$) and (0.25, $p. = 0.009$) were obtained when the Colors variable was correlated with Gender and SES. A negative correlation between Ethnicity and Colors was insignificant (-0.0488, p. = 0.626). When the effects of SES, gender, ethnicity, Word Analysis, and Vocabulary were partialled out, the Colors variable accounted for 1.78% (p. = 0.192) of the variance in Comprehension.

The Cubes variable was the only Gesell subtest which contributed significantly to any reading subskill with the effects of demographics and the other reading skills partialled out. When the effects of the demographic variables were partialled out, the Cubes subtest influenced Reading 3 and Reading 5 (0.3562, p. = 0.001 and 0.2789, p. = 0.008) significantly.

Gender differences were observed with females having better overall scores on most DIAL-R and Preschool Screening Battery subtests than males. These findings were similar to those found by Sears and Keogh (1993). In contrast, males scored higher than females on the Gesell Maturity variable. On the Preschool Screening Battery, females scored significantly higher than males on Commissions, Colors, and Auditory Memory.

<u>Early Determinants Effect on Different Outcome Measures</u>

The same predictors which influenced group membership in the predictive discriminant analysis also emerged as important variables in the multiple regression analysis. Concept development, fine motor skills, and expressive language skills were the most important contributors to three outcome variables, Word Analysis, Vocabulary, and Comprehension.

Visual motor integration, as measured by the Developmental Test of Visual Motor Integration (DTVMI), was significantly associated with Word Analysis at the third grade. The DTVMI was not used in predictive discriminant analyses because it did not improve the discriminatory power of the Preschool Screening Battery at third or fifth grades. It did, however, influence Word Analysis, a finding which supported the hypothesis that skills measured on the DTVMI are more influential in the earlier stages of reading acquisition (Flynn & Flynn, 1978; Klein, 1978; Roth et al., 1993; Vacc et al., 1987) than later when comprehension skills are emerging.

Although the PPVT-R had a relatively high correlation with the EOWPVT (0.63), the two language tests performed quite differently when associated with Comprehension and Word Analysis in multiple regression analyses. The PPVT-R did not contribute significantly to either Comprehension or Word Analysis while the EOWPVT influenced both significantly. When the effects of demographics and other reading subskills were partialled out, the PPVT-R and EOWPVT both contributed significantly to the Vocabulary subtest, but they did not contribute to Word Analysis or Comprehension.

Classification Errors

Significant differences emerged from comparisons of subjects in the 1 to 3 and 3 to 1 groups (subjects with at least 1 group differential between actual and predicted group membership) to other sample subjects. Gesell subjects in the 1 to 3 group had predicted Group 3 membership, yet had actual Group 1 membership. The number of subjects retained in the 1 to 3 group (66%) was proportionally larger than the total number of Gesell subjects retained (32%). However, allowing these subjects to remain at home for an additional year or to spend two years in kindergarten did not appear to improve their reading achievement which was in the below average reading group at third grade. The findings of this study suggested that the practice of providing a catch up year for children who were determined to be developmentally young by Gesell screening did not improve the reading achievement of the retained Gesell subjects. The results do not support the work of Ames and Ilg (1964) and Wood et al. (1984) who reported significant results when school readiness, as measured on the Gesell and other tests, was used to predict school success at the end of kindergarten and sixth grade.

Other factors which seemed to affect classification errors were individual determinants such as chronic illness or behavioral problems. Intra-subtest scatter also seemed to affect classification accuracy.

Implications of the Study

1. The predictive validity of the Preschool Screening Battery (PSB), as estimated by this study, was of sufficient strength to support its use as a preschool screening device. The Language variable which consisted of the Expressive One Word Picture Vocabulary Test (EOWPVT) and the Peabody Picture Vocabulary Test-Revised (PPVT-R) was the single most important predictor of reading group membership at both grade levels. The second most important Preschool Screening Battery predictor was the Cognition variable. The Cognition variable was similar to the Concepts variable on the DIAL-R which was the most significant contributor on that battery. The Preschool Screening Battery had the highest classification ratios for overall and separate group results with the exception of the DIAL-R Group 1 predictions at the third grade.

2. The DIAL-R discriminatory power to predict Group 1 membership was highest (hit ratio = 90%) at the third grade.

This power dropped to 56.30% at the fifth grade. The discriminatory power at the third grade was quite high suggesting that an accurate classification rule for Group 1 membership may be derived from the discriminatory scores obtained from this analysis. The popularity of the DIAL-R (in North Carolina and in the nation), makes this a very promising possibility.

3. The predictive validity of the Gesell School Readiness Test, as estimated by this study, was not significantly better than chance in most cases. The maturationist point of view regarding school readiness attributed to Ilg and Ames (1965) was not corroborated by these findings. The behavioral maturity, purportedly measured by the Gesell, did not accurately predict readiness to read as measured by predicted reading group membership at third and fifth grades. The discriminatory power of the Gesell School Readiness Test to predict reading group membership was low compared to ratios obtained from the other two batteries, and the Gesell did not attain better than chance results (except when demographics were included in the analysis at the fifth grade). Based on these findings, the Gesell School Readiness Test would not be recommended for use in predicting later reading achievement

4. The findings of this study indicate that the predictive validity of the Preschool Screening Battery and

DIAL-R justify their continued use in screening programs. Furthermore, their discriminatory power to correctly classify subjects as below average readers justifies further cross validation using the discriminant scores obtained in this research to identify children at-risk for reading problems in the future. The cost of adding this component to an existing screening program would be minimal. The additional costs would be incurred in providing specific intervention programs for these children.

5. The practice of placing children (designated as developmentally young on Gesell screening) in transitional kindergarten or transitional kindergarten-first grade classes did not affect reading group membership positively at the third grade for four subjects in this study. In fact, the early retentions appeared to impact negatively on the students' achievement. These findings support the work of May and Welch (1984) who found that the practice of buying an extra year for children, based on Gesell screening procedures at kindergarten entrance, did not improve their standardized achievement scores at 2nd, 3rd, 4th, and 6th grades.

6. The findings of this study reinforce the need for using caution in the interpretation of individual screening scores. Any factors which might inhibit optimal performance, such as health, handicapping conditions, or

lack of experiences should be considered before decisions are made regarding the appropriateness of intervention for an individual child.

7. This research supported the multiple-cause theories of reading discussed by Stanovich et al. (1984) in which early reading acquisition is purported to be dependent upon several independent skills which become interrelated at higher levels of reading skill. Although the patterns of influence were consistent from third to fifth grade, in most cases, the influence of the constructs varied when associated with different reading subskills. A developmental trend was noted in which certain constructs were associated with beginning reading skills and others were associated with more advanced skills. For example, Word Analysis was influenced by knowledge and understanding of concepts, fine motor skills related to eye-hand coordination, and visual perception. Vocabulary skills were influenced by conceptual development, and expressive and receptive language skills, eye hand coordination, and abstract thinking skills. Comprehension was influenced by language, concepts, abstract thinking, and auditory memory when the effects of demographics were partialled out. However, when the effects of Word Analysis and Vocabulary were also partialled out, none of the predictor variables

with the exception of the Cubes subtest influenced Comprehension significantly.

Recommendations for Future Research

1. Additional cross validation studies are recommended to assess the classification accuracies of the DIAL-R and the Preschool Screening Battery with other samples. The ultimate outcome of such research would be the development of a classification rule for use in identifying kindergarten students at-risk for reading disabilities.

2. This study focuses on the predictive validity of the three preschool screening batteries. Small n:p ratios precluded more indepth analyses of the relationship between predictors, demographic variables, and outcome measures. The same data analyzed with path analysis might provide more information regarding the interrelationships between demographic variables, predictor variables, intervening reading subskills, and reading achievement at the fifth grade.

3. One limitation of this study is that individual item responses were not available for DIAL-R subjects which prevented factor analysis of the DIAL-R data. Because of this, the DIAL-R subtests were presumed to measure the constructs suggested by their names. Qualitative analysis of the subtests revealed similarities between DIAL-R items

and those on the Preschool Screening Battery and the Gesell School Readiness Test. However, items such as block building and visual discrimination tasks, included on the Motor subtest of the DIAL-R, loaded heavily on the PSB Cognition subtest. This suggested that items included within DIAL-R subtests are not the same as those included in like-named variables extracted from the Preschool Screening Battery and Gesell School Readiness Test factor analysis. Factor analysis of individual DIAL-R items is desirable in order to determine if the PSB and GSRT factor analyses are replicable with DIAL-R data. If the results are replicated, the hypothesis regarding underlying constructs postulated these findings would be strengthened.

4. Additional research is recommended which focuses on the influence of these predictor variables on early reading acquisition (specifically at the first and second grade). For example, the effect of the DTVMI on early acquisition is unclear. Despite the findings of Roth et al. (1993), in which fine motor skills were significant predictors of first grade reading achievement, the DTVMI did not significantly influence total reading achievement at third or fifth grade in this study. It did, however, influence Word Analysis, a subskill, at the third grade. Additional predictive discriminant analyses using the screening batteries to predict reading achievement at the first and second grades

would be useful in testing the hypothesis that determinants such as fine motor skills and other critical subskills exert more independent influence at the beginning stages of reading acquisition than at later stages.

Summary

Predictive validity analyses yielded significant results for two of the preschool screening batteries included in this study, the Preschool Screening Battery and the DIAL-R. Underlying constructs extracted from the three batteries were found to influence reading acquisition over time. These findings were summarized and interpreted as they related to the research questions. Implications were developed for the use of two of the batteries in future preschool screening programs and for the interpretation of the underlying constructs. This chapter was concluded with a recommendations for further research.

BIBLIOGRAPHY

Abrahamson, D. S. & Bell, A. E. (1979). Assessment of the school readiness section of the Early Detection Inventory: Preschool prediction across situational factors. Journal of School Psychology, 17, 162-171.

Aslin, R. N. & Smith, L. B. (1988). Perceptual development. Annual Review of Psychology, 39, 435-473.

Airasian, P. W. (1989). Review of the California Achievement Tests, Forms E and F. In J. C. Conoley, J. J. Kramer, & L. L. Murphy (Eds.), The tenth mental measurements yearbook. Lincoln, NE: The University of Nebraska Press.

Ames, L. B. & Ilg, F. L. (1964). Gesell behavior tests as predictive of later grade placement. Perceptual and Motor Skills, 19, 719-722.

Ary, D., Jacobs, L. C. & Razavieh, A. (1972). Introduction to research in education. New York: Holt, Rinehart, and Winston.

Barnett, D. W., Faust, J. & Sarmir, M. A. (1988). A validity study of two preschool screening instruments: The LAP-D and DIAL-R. Contemporary Educational Psychology, 13, 26-32.

Beery, K. E. (1989). Revised administration, scoring, and teaching manual for the Developmental Test of Visual

Motor Integration. Cleveland, OH: Modern Curriculum Press.

Blau, P. M. & Duncan, O. D. (1967). The American occupational structure. New York: Wiley.

Bloom, B. S. (1964). Stability and change in human characteristics. New York: Wiley.

Bradley, R. H. (1985). Review of the Gesell School Readiness Test. In J. V. Mitchell (Ed.), The ninth mental measurements yearbook. Lincoln, NE: The University of Nebraska Press.

Bryant, D. M., Clifford, R. M. & Peisner, E. S. (1989). Best practices for beginners: Quality programs for kindergartners. Chapel Hill, NC: University of North Carolina at Chapel Hill, Frank Porter Graham Child Development Center.

Bruck, M. (1985). The adult functioning of children with specific learning disabilities: A follow-up study. In I. E. Seigel (Ed.), Advances in applied developmental psychology (pp. 91-126). Norwood, NJ: Ablex.

Butler, S. R., Marsh, J. W., Sheppard, M. J. & Sheppard, J. L. (1985). Seven-year longitudinal study of the early prediction of reading achievement. Journal of Educational Psychology, 77, 349-361.

Caston, R. J. (1989). Dimensions of occupational inequality and Duncan's socioeconomic index. Sociological Forum, 4(3), 329-347.

Colarusso, R., Gill, S., Plenkenhorn, A. & Brooks, R. (1980). Predicting first-grade achievement through formal testing of 5-year-old high-risk children. Journal of Special Education, 14, 355-361.

Cummings, J. A. (1985). Review of Expressive One-Word Picture Vocabulary Test. In J. C. Conoley, J. J. Kramer, & L. L. Murphy (Eds.), The ninth mental measurements yearbook. Lincoln, NE: The University of Nebraska Press.

Duffey, J. B., Ritter, D. R. & Fedner, M. (1976). Developmental Test of Visual-Motor Integration and the Goodenough Draw-A-Man Test as predictors of academic success. Perceptual Motor Skills, 43, 543-546.

Duncan, O. D. (1961). A socioeconomic index for all occupations. In A. J. Reiss (Ed.), Occupations and social status (pp. 109-138). New York: The Free Press of Glencoe, Inc.

Edgington, R. E. (1975). SLC children: A 10-year follow-up. Academic Therapy, I, 53-64.

Fletcher, J. M. & Satz, P. (1980). Developmental changes in the neuropsychological correlates of reading achievement: A six-year longitudinal followup. Journal of Clinical Neuropsychology, 2(1), 23-37.

Fletcher, J. M., Satz, P. & Morris, R. (1984). The Florida Longitudinal Project: An historical review. In S. A. Mednick & M. A. Harway (Eds.), Handbook of Longitudinal Research.

Flynn, T. M. & Flynn, L. A. (1978). Evaluation of the predictive validity of five screening measures administered during kindergarten. Journal of Experimental Education, 46, 65-69.

Gibson, E. J. (1969). Principles of perceptual learning and development. New York: Appleton-Century-Crofts.

Gowen, J. (1989). A new day for young children with handicaps and their families: Report to the Governor. Raleigh, NC: Interagency Coordinating Council.

Hagin, R. (1984). Effects on first-grade promotion practices of a program for the prevention of learning disabilities. Psychology in the Schools, 3, 471-476.

Hair, J. F., Jr., Anderson, R. E., Tathuma, R. L. & Black, W. C. (1992). Multivariate analysis with readings. New York: MacMillian.

Hauser, R. M. & Featherman, D. L. (1977). The process of stratification. New York: Academic Press.

Huberty, C. J. (1994). Applied discriminant analysis. New York: John Wiley & Sons, Inc.

Ilg, F. L. & Ames, L. B. (1965). School readiness. New York: Harper & Row.

Jensen, R. A. (1980). Bias in mental testing. New York: The Free Press.

Johnson, J. H. & Goldman, J. (Eds.). (1990). Developmental assessment in clinical child psychology. New York: Pergamon Press.

Kaufman, A. S. (1971). Piaget and Gesell: A psychometric analysis of tests built from their tasks. Child Development, 42(5), 1341-1360.

Klein, A. E. (1978). The validity of the Beery Test of Visual-Motor Integration in predicting achievement in kindergarten, first and second grades. Educational and Psychological Measurement, 38, 457-461.

Kochanek, T. T., Kabacoff, R. I. & Lipsitt, L. P. (1990). Early identification of developmentally disabled and at-risk preschool children. Exceptional Children, 56(6), 528-538.

Lewis, A. (1980). The early identification of children with learning difficulties, 13, 102-108.

Lichtenstein, R. (1981). Comparison of two preschool screening tests. Journal of Learning Disabilities, 14, 68-72.

Linquist, G. T. (1987). Preschool screening as a means of predicting later reading achievement. Journal of Learning Disabilities, 15, 331-332.

May, D. C. & Welch, E. L. (1984). The effects of developmental placement and early retention on children's later scores on standardized tests. Psychology in the Schools, 21, 381-385.

McCallum, S. (1985). Review of the Peabody Picture Vocabulary Revised. In J. V. Mitchell, Jr. (Ed.), The ninth mental measurements yearbook. Lincoln, NE: The University of Nebraska Press.

Meisels, S. J. (1987). Uses and abuses of developmental screening and school readiness testing. Young Children, 42(2), 4-6, 69-73.

Meisels, S. J., Wiske, M. S. & Tivnan, T. (1984). Predicting school performance with the Early Screening Inventory. Psychology in the Schools, 21, 25-33.

Miller, L. J. & Sprong, T. A. (1986). Psychometric and qualitative comparison of four preschool screening instruments. Journal of Learning Disabilities, 19(8), 480-484.

Moore, R. & Zuckerman, M. (1992). Report of student performance, Spring, 1992. Raleigh, NC: Division of Accountability Services, North Carolina Department of Public Instruction.

Muehl, S. & Forell, E. R. (1973). A follow-up study of disabled readers: Variables related to high school reading performance. Reading Research Quarterly, 9, 110-123.

North Carolina Department of Public Instruction. (1993). 1993-94 North Carolina State Testing Results. Raleigh, NC.

Norusis, J. (1993). SPSS for Windows Base System User's Guide Release 6.0 {Computer Program Manual}. Chicago, IL: SPSS, Inc.

Pommerich, M., Belleaud, K., Williams, V. & Thissen, D. (1993). User's Guide for the North Carolina End of Grade Tests. Chapel Hill, NC: Psychometric Laboratory, University of North Carolina.

Poteat, G. M. (1989). Review of the Developmental Indicators for the Assessment of Learning-Revised. In J. C. Conoley, J. J. Kramer & L. L. Murphy (Eds.), The tenth mental measurements yearbook. Lincoln, NE: The University of Nebraska Press.

Rafoth, M. A. (1988). Early identification of learning disabilities using the Meeting Street School Screening Test. Journal of Learning Disabilities, 21(3), 186-188.

Reynolds, C. R., Wright, D. & Wilkinson, W. A. (1980). Incremental validity of the test for auditory comprehension of language and the developmental test of visual-motor integration. Educational and Psychological Measurement, 40, 503-507.

Reiss, A. J. (Ed.). (1961). Occupations and social status. New York: The Free Press of Glencoe, Inc.

Roth, M., McCaul, E. & Barnes, K. (1993). Who becomes an "at-risk" student? The predictive value of a kindergarten screening battery. Exceptional Children, 59(4), 348-358.

Salvia, J. & Yesseldyke, J. (1985). Assessment in special and remedial education, (3rd ed.) Boston: Houghton Mifflin.

Sattler, J. (1992). Assessment of children (rev. 3rd ed.). San Diego: Jerome M. Sattler.

Satz, P. & Fletcher, J. M. (1988). Early identification of learning disabled children: An old problem revisited. Journal of Consulting and Clinical Psychology, 56(6), 824-829.

Sears, S. & Keogh, B. (1993). Predicting Reading Performance Using the Slingerland Procedures. Annals of Dyslexia, 43, 78-89.

Sedlmier, P. & Gigerenzer, G. Do studies of statistical power have an effect on the power of studies? Psychological Bulletin, 105, 309-316.

Snider, V. E. & Tarver, S. (1987). The effect of early reading failure on acquisition of knowledge among students with learning disabilities. Journal of Learning Disabilities, 20(6), 351-356.

Spivack, G. M. (1985). Review of Expressive One Word Picture Vocabulary Test. In J. C. Conoley, J. J. Kramer & L. L. Murphy (Eds.), The ninth mental measurements yearbook. Lincoln, NE: The University of Nebraska Press.

SPSS, Inc. (1994). SPSS for Windows Update {Computer Program Manual}. Chicago, IL: SPSS, Inc.

SPSS, Inc. (1995). SPSS for Windows Advanced Statistics V6.1 {Computer Program}. Chicago, IL: SPSS, Inc. SPSS, Inc.

Stallings, J. A. & Stipek, D. (1986). Research on early childhood and elementary school teaching programs. In M. C. Wittrock (Ed..), Handbook on research on teaching (3rd ed., pp. 727-751).

Stanovich, K. E., Cunningham, A. E. & Feeman, D. J. (1984). Intelligence, cognitive skills, and early reading progress. Reading Research Quarterly, 19(3), 278-298.

Stevens, J. (1992). Applied multivariate statistics for the social sciences. Hillsdale, NJ: Lawrence Erlbaum Associates.

Stevens, G. & Cho, J. H. (1985). Socioeconomic indexes and the new 1980 Census occupational classification scheme. Social Science Research, 14, 142-168.

Stevenson, H. W. & Newman, R. S. (1986). Long-term prediction of achievement and attitudes in mathematics and reading. Child Development, 57, 646-659.

Stricker, L. J. (1988). Measuring social status with occupational information: A simple method. Journal of Applied Social Psychology, 18(5), 423-437.

Suen, J. K., Czudnowski, C. M. & Goldenberg, D. S. (1989). Classification reliability of the Dial-R Preschool Screening Test. Educational and Psychological Measurement, 49(3), 673-680.

Teeter, P. A. (1985). Neurodevelopmental investigation of academic achievement: A report of years 1 and 2 of a longitudinal study. Journal of Consulting and Clinical Psychology, 53, 709-717.

Telegdy, G. A. (1975). The effectiveness of four readiness tests as predictors of first grade academic achievement. Psychology in the School, 12, 4-11.

Tramontana, J. G., Hooper, S. R. & Selzer, S. C. (1988). Research on the preschool prediction of later academic achievement: A review. Developmental Review, 8, 89-146.

Vacc, N. A., Vacc, N. N. & Fogleman, M. S. (1987). Preschool screening: Using the DIAL as a predictor of first-grade performance. Journal of School Psychology, 25(1), 45-51.

Wardrop, J. L. (1989). Review of the California Achievement Tests, Forms E and F. In J. C. Conoley, J. J. Kramer, & L. L. Murphy (Eds.), The tenth mental measurements yearbook. Lincoln, NE: The University of Nebraska Press.

Waters, E. (1985). Review of the Gesell School Readiness Test. In J. V. Mitchell, Jr. (Ed.), The ninth mental measurements yearbook (pp. 610-611). Lincoln, NE: The University of Nebraska Press.

Wood, C., Powell, S. & Knight. (1984). Predicting school readiness and the validity of developmental age. Journal of Learning Disabilities, 17(1), 6-11.

Yesseldyke, J. E. & O'Sullivan, P. J. (1987). Predicting preschool screening referral rates from district demographic data. Journal of School Psychology, 25, 119-129.

Appendix A

Letters to Superintendents

Sandra Fortenberry Dill
201 Beech Tree Circle
Mount Airy, NC 27030

Licensed Psychological Associate

Telephone (910)-786-9521

Dear Superintendent:

I am a graduate student at Virginia Tech University in Blacksburg, Virginia completing requirements for a doctorate in school psychology and counseling. For my dissertation, I am conducting predictive validity studies on three preschool screening batteries used in North Carolina during the 1988-89 school year.

The Developmental Indicators for the Assessment of Learning-Revised (DIAL-R), which was used in your school system for several years, is one of these batteries. I some of your staff members last week regarding the availability of preschool screening data, and the appropriate procedure for requesting approval to collect data in your school system.

I understand preschool screening data are stored in the central office, and would be available without any preparation time from staff members. I would collect the remaining data from student records after obtaining informed consent forms from the parents.

A copy of the parent information letter, the consent form, and a project abstract are attached. Subjects would be recruited from upcoming sixth graders in August or early September of the 1994-1995 school year.

Thank you in advance for your time and consideration.

Sincerely,

Sandra F. Dill

Sandra Fortenberry Dill
201 Beech Tree Circle
Mount Airy, NC 27030

Licensed Psychological Associate

Telephone (910)-786-9521

Dear Superintendent:

I am a graduate student at Virginia Tech University in Blacksburg, Virginia completing requirements for a doctorate in school psychology and counseling. For my dissertation, I am conducting predictive validity studies on three preschool screening batteries used in North Carolina during the 1988-89 school year.

The Gesell School Readiness Test which was used in your school system for several years, is one of these batteries. I some of your staff members last week regarding the availability of preschool screening data, and the appropriate procedure for requesting approval to collect data in your school system.

I understand preschool screening information is filed in the student cumulative folders, and would be available without any preparation time from staff members. I would collect the data from student records after obtaining informed consent forms from the parents.

A copy of the parent information letter, the consent form, and a project abstract are attached. Subjects would be recruited from upcoming sixth graders in August or early September of the 1994-1995 school year.

Thank you in advance for your time and consideration.

Sincerely,

Sandra F. Dill

Sandra Fortenberry Dill
201 Beech Tree Circle
Mount Airy, NC 27030

Licensed Psychological Associate Telephone (910)-786-9521

Dear Superintendent:

I am a graduate student at Virginia Tech University in Blacksburg, Virginia completing requirements for a doctorate in school psychology and counseling. For my dissertation, I am conducting predictive validity studies on three preschool screening batteries used in North Carolina during the 1988-89 school year.

The Preschool Screening Battery, which was used in your school system for several years, is one of these batteries. I some of your staff members last week regarding the availability of preschool screening data, and the appropriate procedure for requesting approval to collect data in your school system.

I understand preschool screening information is filed in the student cumulative folders, and would be available without any preparation time from staff members. I would collect the data from student records after obtaining informed consent forms from the parents.

A copy of the parent information letter, the consent form, and a project abstract are attached. Subjects would be recruited from upcoming sixth graders in August or early September of the 1994-1995 school year.

Thank you in advance for your time and consideration.

Sincerely,

Sandra F. Dill

Appendix B

Parent Letters and Informed Consent Forms

141 Jackson Road
Mount Airy, NC 27030

Dear Parents:

I am a graduate student at Virginia Polytechnic Institute and State University in Blacksburg, Virginia, conducting research on how a child's development at school entrance is related to reading achievement at the sixth grade.

This study has been approved by your school principal and your school system. If you give approval, the following will be collected from your child's school records: (1) preschool screening scores, (2) achievement test scores, (3) number of retentions, (4) reading grades, and (5) placement in Exceptional Children Programs.

I assure you that all information collected will remain confidential as each student will be assigned an identification number. Data collected will be identified by student number alone, and will not be seen by anyone other than the researcher.

Your participation is completely voluntary and requires that you sign and return the consent form. Nothing else will be required of you, and your child will not be affected in any way by this project.

Please read and return the signed consent form to your child's teacher within five school days. If you have questions, please contact your child's principal or contact me at 141 Jackson Road, Mount Airy, North Carolina 27030.

Thank you in advance. I hope you will agree to help with this important study by allowing your child's records to be reviewed. Results of study will be available at your child's school upon request.

Thank You,

Sandra Dill

Sandra Fortenberry Dill
141 Jackson Road
Mt. Airy, NC 27030

March 16, 1995

Dear Parents or Guardians:

A few weeks ago you received a letter requesting permission to for me to obtain information from your child's school records for a study on preschool screening and how it predicts reading achievement.

If you have not returned the consent form at this time, I hope you will reconsider. Your assistance in this research would be greatly appreciated. Again, this would not take any additional effort or time from you or your child, and I want to assure you all information will be handled confidentially as the enclosed letter indicates.

Again, if you have any questions about this study, you can contact me at 141 Jackson Road, Mt. Airy, North Carolina.

Your help is greatly needed. I hope to hear from you soon.

Sincerely,

Sandra Dill

PREDICTIVE VALIDITY STUDIES OF
THREE PRESCHOOL SCREENING BATTERIES
Virginia Polytechnic Institute and State University
Informed Consent For Release of Information

Student Name________________________Researcher: Sandra Dill

Mother's Occupation____
Job Title or if self-employed type of business
Father's Occupation____
Job Title or if self-employed type of business

I give consent for information on my child to be included in this study. I understand this study involves a review of each child's school records which may include: (1) the cumulative record, and (2) exceptional children files for the purpose of collecting the following:

Kindergarten Screening Scores	**Achievement Test Scores**
Grade Retentions	**Reading Grades**
Exceptional Children Placement Data	**Student Background Data**

I understand information collected will remain confidential, and will not be seen by anyone other than the researcher.

****

Subjects Approval

_____I do give permission for my child's records to be reviewed. I have read and understand the informed consent, and I agree to the release of the information described above to the researcher for the purposes of this study.

_____I do not give permission for my child's records to be reviewed. I have read and understand the informed consent, and I do not agree to the release of the information described above to the researcher for the purposes of this study.

______________________ ____________________ _________
Parent/Legal Guardian Relation to student Date

Appendix C

Data Collection Forms

DIAL-R

DATA COLLECTION

Subject____________________________DOB________________DOE__________

Race__________SEX_________CA__________DIAL-R TOTAL_______________

DIAL-R

MOTOR____________ CONCEPTS_________ LANGUAGE______________

CALIFORNIA ACHIEVEMENT TEST

	SCALE	GE	%	NCE
WORD ANALYSIS				
VOCABULARY				
COMPREHENSION				
TOTAL TEST				

END OF GRADE TESTS

	DS	%
4TH	_______	_______
5TH	_______	_______

RETENTIONS_______________ EC PLACEMENT__________________________

GRADES

_________1ST _________2ND _________3RD _________4TH ________5TH

MOTHER'S OCCUPATION____

FATHER'S OCCUPATION____

COMMENTS____

____

GESELL SCHOOL READINESS TEST
DATA COLLECTION

NUMBER__________________________DOB__________DOE__________

RACE________ DA____________________GENDER____________

C-A AT SCREENING_____________

CUBES TESTS__________________

COPYING FORMS________________

INCOMPLETE MAN_______________

WRITING NAME/NUMBERS _________

NAMING ANIMALS/INTERESTS_______

INTERVIEW____________________

****

CALIFORNIA ACHIEVEMENT TEST

	SCALE	GE	%	NCE
WORD ANALYSIS	____	____	____	____
VOCABULARY	____	____	____	____
COMPREHENSION	____	____	____	____
TOTAL TEST	____	____	____	____

END OF GRADE TESTS

	DS	%
4TH	_____	_____
5TH	_____	_____

RETENTIONS________EC PLACEMENT_________________________

GRADES

______1ST ______2ND ______3RD ______4TH ______5TH

MOTHER'S OCCUPATION SEI______

FATHER'S OCCUPATION SEI_______

COMMENTS____

____

PRESCHOOL SCREENING BATTERY

DATA COLLECTION FORM

SUBJECT NUMBER________________DOB__________________DOE____________

RACE____________________ SEX______________ C-A__________________

LEARNING INSTRUMENT________TT SCORE PPVT-R____________________

SS % RS AE

BLOCK BUILDING__________________

DMTVI-R______________

COLORS__________________________ RS % SS AE

VISUAL MEMORY___________________ MOTOR_________________

TT AE

VISUAL DISCRIMINATION__________

EOWVT_________________

AUDITORY MEMORY_________________ RS AE

COMMISSIONS_____________________

BEHAVIORAL

GEN INFOR________________________ ______________________

CONCEPTS_________________________ LD____________________

DOP TYPE

NUMBERS__________________________

MOTHER'S OCCUPATION____ _______

RATING________

FATHER'S OCCUPATION_____________________________________ _______ CS

RATING

CALIFORNIA ACHIEVEMENT TEST

	SCALE	GE	%	NCE
WORD ANALYSIS				
VOCABULARY				
COMPREHENSION				
TOTAL TEST				

TYPE OF TEXT

____

RETENTIONS

____

4TH END OF GRADE READING_______________

5TH END OF GRADE READING_______________

GRADES

______1ST ______2ND ______3RD ______4TH ______5TH

Appendix D

Typicality Tables

Table D-1

Typicality Differences DIAL-R Group

Group	Subtest	t-test	p-value
1	Motor	2.19	0.034
	Concepts	9.14	0.000
	Language	4.90	0.000

Table D-2

Typicality Differences GSRT Group

Group	Subtest	t-value	p-value
1	Writing Name/Numbers	5.57	0.000
	Naming Animal and Interests	8.28	0.000
2	Cubes	4.13	0.000
	Incomplete Man	3.80	0.001
	Naming Animals/Interests	-4.75	0.000
	Interview	3.08	0.004
3	Cubes	-4.08	0.000
	Copyform	-4.31	0.000
	Writing Name and Numbers	-3.56	0.001
	Interview	-5.81	0.000

Appendix E

Comparison of Means Tables

Table E

Comparison of Means between Group 1 to 3 Gesell and DIAL-R Subjects and other Gesell and DIAL-R Group 3 Subjects

Battery	Subtest	t-value	p-value
Gesell	Cubes	3.55	0.001*
	Copy Forms	1.53	0.135
	Incomplete Man	1.88	0.070
	Naming Animals/Interests	3.11	0.004
	Writing Name/Numbers	2.19	0.036
	Interview	-1.92	0.064
DIAL-R	Concepts	-1.49	0.146
	Motor	-3.20	0.003
	Language	0.44	0.666

Appendix F

Partial Correlations Tables

Table F-1

Partial Correlations Between DIAL-R Variables and Reading Subskills Controlling for Demographics

Subtest	Word Analysis	Vocabulary	Comprehension
Concepts	0.4620**	0.5177**	0.4522**
Motor	0.3874**	0.3662**	0.2957**
Language	0.3636	0.3482**	0.3211**

* p. = .05, ** p. = .01

Table F-2

Partial Correlations Between Gesell Variables and Reading Subskills Controlling for Demographics

Subtest	Word Analysis	Vocabulary	Comprehension
Copy Forms	0.3711**	0.3728**	0.2575*
Cubes	0.4305**	0.3593**	0.1939
Inc. Man	0.4848**	0.3955**	0.3497**
Interview	0.2744*	0.2447*	0.2356
Writ/Name/Num	0.2721*	0.3129*	0.2078
Naming	0.1840	0.2169	0.1436

* p = .05, ** p. = .01

Table F-3

Partial Correlations Between PSB Variables and Reading Subskills Controlling for Demographics

Subtest	Word Analysis	Vocabulary	Comprehension
Colors	0.0805	0.1573	0.1977*
DTVMI	0.2628**	0.2254*	0.2416**
EOWPVT	0.2801**	0.4319**	0.3613**
PPVT-R	0.1945*	0.3974**	0.3013**
Cognition	0.2556*	0.3127**	0.2772**
Perception	0.2407*	0.2478**	0.2928**

* p. = .05, p. = .01

Table F-4

<u>Partial Correlations Between Predictor Variables and Reading Subskills Controlling for Demographics and Other Reading Subskills</u>

Subtest	Word Analysis	Vocabulary	Comprehension
DIAL-R			
Concepts	0.1016 (0.332)	0.2106 (0.043)*	0.0911 (0.385)
Motor	0.1764 (0.091)	0.1049 (0.317)	-0.0033 (0.975)
Language	0.1449 (0.166)	0.0703 (0.503)	0.0651 (0.535)
GSRT			
Copy Forms	0.2003 (0.107)	0.2154 (0.082)	-0.1286 (0.304)
Cubes	0.3374 (0.006)	0.2425 (0.050)*	-0.2644 (0.032)*
Incomplete Man	0.1901 (0.126)	0.1365 (0.275)	0.0039 (0.975)
Naming	0.0668 (0.594)	0.1408 (0.259)	-0.0676 (0.590)
Writing Name/Numb	0.1089 (0.0668)	0.2035 (0.101)	-0.1021 (0.415)
PSB			
Colors	-0.0683 (0.506)	0.0219 (0.832)	0.1336 (0.192)
DTVMI	0.1330 (0.194)	0.0019 (0.985)	0.0721 (0.482)
PPVT-R	-0.1135 (0.268)	0.2947 (0.003)*	-0.0076 (0.941)
EOWPVT	-0.0324 (0.752)	0.2495 (0.014)*	0.0311 (0.754)
Cognitive	0.0501 (0.626)	0.1291 (0.207)	0.0328 (0.750)
Perception	0.0656 (0.523)	-0.0038 (0.970)	0.1441 (0.159)

Appendix G

Group Centroids

Table G

Group Centroids

Group	Third Grade		Fifth Grade	
	Function 1	Function 2	Function 1	Function 2
DIAL-R				
1	-1.15943	0.04310	-0.48480	-0.29091
2	0.49108	-0.21656	-0.17186	0.37242
3	1.24216	0.13891	0.70845	-0.10988
GSRT				
1	-0.75238	-0.20033	-0.80188	0.06040
2	-0.13076	0.26698	-0.14866	0.07489
3	0.70611	-0.11379	0.71469	0.03226
PSB				
1	-1.11061	-0.29793	1.14238	-0.28160
2	-0.37572	0.21380	-0.47988	0.29829
3	0.66765	-0.06995	1.06411	-0.06905

SANDRA FORTENBERRY DILL

VITA

Sandra Fortenberry Dill was born in Leland, Mississippi and attended public schools in Leland, Mississippi, Mesilla Park, New Mexico, and Alexandria City Schools, Alexandria, Virginia. The author received a B. S. degree from Indiana State University In English and secondary education in 1970 and received an M. S. degree in Psychology from Radford University, Radford, Virginia in 1972.

The author taught English at Lancaster High School, Kilmarnock, Virginia during 1971 and 1972 before entering the School Psychology Program at Radford University in 1972. She was employed as a school psychologist for three years in Alamance County Schools, Graham, North Carolina before becoming Director of Exceptional Children, a position she held for 14 years.

In 1991, she became a full-time doctoral student at Virginia Tech in Student Personnel Services, and received the Doctor of Philosophy degree with a specialization in counseling and school psychology in 1995. She completed a six month internship in a mental health setting and has been employed in private practice. She is licensed as a psychological associate by the North Carolina State Board of Examiners.

The author is currently employed as a staff psychologist in Guilford County Schools, Greensboro, North Carolina. She is certified by the North Carolina State Department of Public Instruction as a school psychologist and program administrator, and holds various other certifications in special education, supervision, and curriculum.

Sandra Fortenberry Dill

Sandra Fortenberry Dill