PREDICTIVE ANALYSIS OF EPSF KINDERGARTEN SCREENING MEASURES TO GATES MACGINITIE READING TESTS PERFORMANCE IN FIRST GRADE STUDENTS

By

GREGORY WAYNE REED
Bachelor of Science
Southwestern Oklahoma State University
Weatherford, Oklahoma
1973

Masters of Education
Southwestern Oklahoma State University
Weatherford, Oklahoma
1974

Submitted to the Faculty of
the Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the degree of
DOCTOR OF PHILOSOPHY
December, 1995

PREDICTIVE ANALYSIS OF EPSF KINDERGARTEN
SCREENING MEASURES TO GATES MACGINITIE
READING TESTS PERFORMANCE IN
FIRST GRADE STUDENTS

Thesis Approved:

Thesis Advisor

Charles R. Davis

Paul G. Ward

Joseph R. Pearl

Ruth E. Jones

Thomas C. Collins

Dean of the Graduate College

ACKNOWLEDGMENTS

I am grateful for this opportunity to formally extend my thanks to all those who have been so supportive and helpful throughout my professional and personal life in the pursuit of this great undertaking. To my mother and brother, I give my thanks for their enduring belief throughout my life in the love of learning and their ever present quest for personal excellence in all tasks attempted.

I wish to thank Dr. Kay Bull, my doctoral committee chairman and advisor, for his helpful, yet sometimes painful, advice that always in the end contributed greatly to the thoroughness and direction of this study. To Dr. Bob Davis and Dr. Joe Pearl many thanks for their willingness to be on my doctoral committee as well as their supportive comments and suggestions. I would especially like to thank Dr. Ruth Tomes for her support and willingness to continue on my doctoral committee at the later stages of my dissertation, despite the potential inconvenience after her change of job location.

My deepest gratitude to Dr. Dale Feuquay, Applied Behavioral Studies (ABSED) Department Head at Oklahoma State University, for those hours of consultation and his patience in helping me through my struggles with statistical theory and its application in my study. Many thanks to Ms. Iris McPherson, for her data analysis expertise and suggestions during the computer analysis phase of my research. My gratitude to Dr. David McIntosh, my former doctoral committee chairman, who helped me during my initial dissertation idea formulations. His love of research and helpful ideas were inspiring.

I would like to now formally recognize Dr. Paul Warden, Director of the Oklahoma State University School Psychology Program, for his many years of leadership in the field of school psychology and his inspiration that lead me to begin my quest at Oklahoma State University in my chosen field initially almost

twenty years ago. How does one thank someone who has been such a major influence in their professional life? Dr. Warden was always available throughout the years with helpful advice and a steadfast belief in my professional potential, even at times when I doubted myself.

To my loving family, how can I find the words to express my gratitude for their love and support during this arduous dissertation and doctoral process. To my loving wife, Debbie, I can only give my love and heartfelt thanks for her tolerance, love, emotional support and efforts in those long hours of reading dissertation copy and helping with all those extra annoying details throughout my "Ph.D years". To my children, those near and those far way, know that I dedicate this effort to you in the belief that through our example, we as parents, help shape the lives of our loved ones.

Dedicated to
Amanda, Stephanie and Chris

TABLE OF CONTENTS

Chapter Page

Chapter Page

LIST OF TABLES

Table Page

Table Page

Chapter 1

Introduction

Background

Early identification of potential learning problems has been a growing area for research and programmatic implementation since the 1960's. Renowned developmental psychologists Jean Piaget and Benjamin Bloom helped raise awareness of the issue of early problem identification in young children by researchers, the general public and governmental agencies throughout the United States.

Developmental psychology's major contribution to the literature is the concept of the significance of preschool years to later subsequent learning. Bryant (1991) stated "Piaget (1952) suggests that a full understanding of human knowledge could be gained through the study of its formation and evolution in childhood", (p. 3). Bloom (1964) commenting on intellectual growth stated '"fifty percent of development takes place between conception and age 4" (p. 88). Bryant (1991) reported that Bruner (1980) commented " the importance of early childhood development to the intellectual, social and emotional growth of human beings is one of the most revolutionary discoveries of modern times" (p. 3).

Federal legislation in the 1960's was instigated in response to increased public awareness generated by the growing body of research on early identification of disabling conditions in infants and preschoolers. The federal government's push for early intervention programs for children resulted in funding of the 1964 Child Health and Mental Retardation Act as well as the Head Start and Follow Through programs. Nuttal, Romero and Kalesnik (1992) attribute the development of Head Start and Follow Through programs as creating "a need for preschool tests for diagnosis, monitoring and program evaluation", (p. 6).

Other significant federal legislation regarding early intervention programs included in 1967 the Early Prevention Screening Diagnosis and Treatment Program (EPSDT) that established Handicapped Children's Early Education Model Programs and Child Service Demonstration Projects (Kelly & Surbeck, 1983). Subsequent significant federal legislation included the Handicapped Children's Early Education Assistance Act of 1968, Public Law 93-380 established Child Find Legislation in 1974, Public Law 94-142 in 1975 required schools to provide intensified services to all severely disabled children below the age of five, and Public Law 99-457 in 1986 required the provision of public school services for disabled children three to five years old and children birth through three in designated high risk developmental areas (Paget, 1990).

State governments have followed the lead of federal mandates through creating required developmental screenings for three to six-year-old children in more than 25 states (Meisels, 1987). Minnesota in 1977 was the first state to have comprehensive free screening to all kindergarten children just two years after the passage of 94-142 (Ysseldyke, Thurlow, O'Sullivan & Bursaw, 1986). Implementation of state and federal legislation for identified at-risk preschoolers generated dramatic yet varied statements regarding the U. S. incidence of at-risk children for potential school failure. These comments include: (1) "9.5 million children or 12% of the school population are impaired by physical, mental or emotional problems" (Liechtenstein and Ireton, 1984, p. 1), (2) Catterall and Cota-Robles (1988) stated "20 million school age children are at-risk of having school problems" (Roth, McCaul, & Barnes,1993, p. 348); (3) Levin (1985) states "one third of U.S. children are educationally at-risk" (Roth, et al. 1993, p. 349). Finally, Olson (1991) stated as many as 40% of U.S. children began school at-risk for school failure.

Advocates for early identification of learning difficulties base their argument on assumptions that: 1) early experience is important to later development, and 2) prevention of predicted learning failure can be done through early intervention. Adelman (1982) effectively sums up this point by stating " prevention and intervention in the earliest stage of a problem seems as having the potential for being more effective and economical than later remediation" (p. 255). Also, 3) young children's behavior is susceptible to change (Mercer, Algozzine & Trifiletti, 1979). Slavin, Karweit, and Wasik (1994) adds 4) that early identification can be more cost effective than later intervention. For example, he states that early intervention can potentially reduce costs of student grade retention in schools up to $5,000 a year per child. Early intervention to increase children's potential for positive change has been widely advocated (Adelman, 1982; Bailey & Wolery, 1989; McGowen, 1991; Ysselkyke, Thurlow, O'Sullivan, & Bursaw, 1986).

Specific benefits of screening preschool children noted by Bailey and Wolery (1989) included individualized program planning, diagnostic placement, program placement, evaluation decisions, and differentiation of at-risk as compared to normal preschool children. Despite the noted needs and benefits of preschool screenings, there exists "no acceptable standards that allow for universal comparison of children's progress" (Mcloughlin & Rausch, 1990). Preschool test and program developers have generated varied materials to meet the demands for effective early identification and intervention with preschool children.

The initial enthusiastic efforts by schools and other social agencies for early identification and remediation of potential learning problems have resulted in some definite criticisms in the research literature. Adelman (1982) comments on widespread application of screening procedures as "another example when pressure and enthusiasm for new procedures have led to inappropriate extrapolation of research findings and premature applications" (p. 255).

Specific criticisms of early childhood screenings include the potential for: (1) premature labeling (Adelman, 1982; Barnes, 1982; Garner, 1993; Lichtenstein & Ireton, 1984; Mercer, Algozzine, & Trifiletti, 1979; Paget & Nagle, 1986; Thurlow, O'Sullivan & Ysseldyke, 1986); (2) limited generalization of results (Lichtenstein & Ireton, 1984; Miller & Sprong, 1986; and McGowen, 1991); (3) difficulty in handling the issue of rapid developmental change by preschool children (Barnes, 1982; Lichtenstein & Ireton, 1984; Zeitlin, 1976) and (4) stress in young children through screening procedures (Elkind, 1989). Mcloughlin and Rausch (1990) stated it is likely that most, if not all, the data obtained from children screenings are underestimates because of their unsophisticated test taking skills, heavy loading of verbal items with higher frequency of undeveloped language skills in children, motivational errors, and test rapport issues with younger children.

Notable early research on screening tests focused on surveys and reports of the psychometric merits of the rapidly generated number of screening instruments credited to potentially assess young children's learning abilities. A 1971 UCLA study of preschool and kindergarten assessment instruments including 120 preschool tests (having 630 total subtests) found "only seven subtests were rated as providing good measurement validity" (Kelley & Surbeck, 1983, p. 12). Joiner (1977) conducted a survey of 177 New York state school districts and found 151 different procedures and/or tests utilized for preschool screening with only 16 having "even marginal reliability and validity".

During the 1980's research efforts continued on the psychometric qualities of screening tests. Meisels (1987) reported a 1984 Michigan Department of Education survey of 111 tests used for preschool, kindergarten, and pre-first grade level programs. "Fewer than ten of these tests were appropriate in terms of age and purpose to which they were put to" (Meisels, 1987, p. 5). Lehr, Ysseldyke and Thurlow (1987) evaluated 109 different preschool tests used by 54 United States Early Education Programs for Handicapped Children (ages birth to six years). Lehr, et al. (1987) found that of the 19 most used tests only three tests had "technically adequate norms, validity and reliability" (p. 397). A 1992 review of eight screening tests found only four were standardized nationally and appropriate for most children (Nuttall, Romero & Kalesnik, 1992).

Adelman (1982) commented that:

> Screening is much in demand for identifying learning problems at an early age. A climate has been established when both consumer and supplier are less critical than they should be in evaluating the validity of proposed and prevalent procedures. (p. 255)

The basic properties and criteria for a good developmental screening test have been discussed at length in the literature including such areas as standardization, reliability, validity, item gradients and test floors (Bracken, 1987). Recommended test-retest and interrater reliability coefficients of .8 are suggested (Lichenstein & Ireton, 1984; Rosenkoetter & Wanska, 1992; Salvia & Ysseldyke, 1991). Predictive validity is one of the major issues in screening tests. Satz and Fletcher (1988) state one of the most frequent problems with preschool screenings is "inadequate assessment of predictive utility of screening devices" (p. 24). The importance of predictive validity issues in preschool screening programs was summarized by Satz and Fletcher (1979), stated "The predictive value of preschool screening programs is directly related to the predictive error rate of the measures employed" (p. 45). Meisels (1989) stated, "developmental screening tests are in widespread use but few reliable and valid tests are available" (p. 578).

The Early Prevention of School Failure (EPSF) is a nationally validated preschool developmental screening program designed to prevent school failure through early identification of four to six year-old children's developmental skills and learning styles (Werner, 1990). EPSF was developed in 1971, nationally validated originally in 1974, again in 1977 with Chapter I and migrant children five foreign countries" (Werner, 1990, p. vi).

EPSF consists of two major program components: 1) a preschool screening battery; and 2) a recommended developmentally appropriate supplemental curriculum for kindergarten and first grade. The first grade curriculum is known as On The Road to Success in Reading and Writing (Success) and was nationally validated by the National Diffusion Network (NDN) in 1990 (EPTW, 1994). This program is designed to provide supplemental intervention to deal with each child's instructional learning modalities as determined from the EPSF screening battery. The Success curriculum focuses on reading and writing skill development and has been used "in 50,000 classrooms over a period of ten years" (Werner, 1991, p. 4). The EPSF program was approved in June 1990 by the U.S. Department of Education as an "exemplary education program" (Betz, 1990). The EPSF Success program is well accepted and widely used by public school systems. However, the initial EPSF screening battery has had very limited predictive validity research in the literature. Only a total of 12 studies on the EPSF program could be found in the literature dating 1984 through 1994. Only four EPSF independent studies involving the predictive validity of the EPSF screening battery were found.

The EPSF screening battery consists of five test instruments including the Peabody Picture Vocabulary Test Revised (PPVT-R), the Preschool Language Scale (EPSF.PLS), the Motor Activity Scale (EPSF.MAS), the Draw-A-Person Test (EPSF.DAP), and the Developmental Test of Visual Motor Integration (VMI) as noted by Werner, 1990. The purpose of the screenings is to "determine the developmental levels of modality skills needed for reading and writing" (Werner 1990, p. i). The tabulated total five EPSF test scores including three MAS and five PLS subtests together generate seven modality area scores. These seven EPSF modality areas are designated: Receptive Language (RL), Expressive Language (EL), Auditory (AU), Visual Memory (VD), Visual Discrimination, Fine (FM) and Gross Motor (GM) skills.

EPSF authors speak of the significance of diagnosing "moderate need" or "considerable need" children. "Moderate need" is defined as one year below chronological age in one of the seven EPSF modality areas. "Considerable need" is defined as two years below chronological age in two of the seven EPSF modality areas (Werner, 1990). A computer generated EPSF screening profile (see Appendix A) compares the preschool child's obtained test results and test norms to determine the supposed significant strengths and/or noted " moderate or high risk" of the child in any of the seven defined modality areas. The profile supposedly is used as an measure of the potential for the individual child's eventual school success or failure (Werner, 1990) .

There has been previous research in the literature on preschool prediction of later academic achievement. Notable examples are Mercer, Algozzine and Trifiletti's (1979) review of 15 studies and Horn and Packard's (1985) meta-analysis of 58 studies in reading from 1960 to 1980. A significant recent meta-analysis by Tramontana, Hooper and Selzer (1988) reviewed a total of 74 studies published from 1973 to 1986 regarding preschool measures and their predictability of later academic achievement, especially reading and math. Tramontana, et al. (1988) found that overall reading prediction had "little agreement among investigators as to the relative effectiveness of cognitive, verbal, and perceptual-motor measures in predicting subsequent reading performance" (p. 101). They found significant predictive relationships when cognitive, verbal and perceptual-motor measures were combined.

The Gates MacGinitie Reading Test is a well known reading achievement measure originally developed in 1926 and was revised most recently for the third time in 1989. The recently revised Gates MacGinitie Reading Test (Third Edition) contains nine levels "to assess student achievement in reading skills from kindergarten through grade 12" (Graham, 1990, p. 21) . The Gates MacGinitie Reading Test yields three scores: Reading Vocabulary, Reading Comprehension and Total Reading. The current study deals with the predictive validity of the EPSF kindergarten screening battery, including its tests, subtests and resultant modality scores, as predictors of later first grade Gates MacGinitie reading achievement.

Significance of Study

A current review of the literature pertaining to the EPSF screening battery finds very limited validity research despite its use with kindergarten and first grade children. Most recently, Gridley, Mucha and Hatfield (1995) in their discussion of preschool screening mentioned 15 "commonly used screening instruments" including the EPSF. Gridley, et al. (1995) mentioned only 3 of the 15 reviewed screening tests and test batteries met all six defined test selection criteria. "Evidence of adequate standardization and psychometric" was one of the six defined test criteria. Only 4 of the 15 reviewed screening tests met this criteria - the EPSF was not one of these.

A total of 12 independent studies with the EPSF program were found in the literature dating from 1984 through 1994, despite its supposed widespread use as an exemplary treatment program. The majority of these studies, noted on page 43 of Chapter II, based their research on the effects of the EPSF treatment program as measured by end of school year achievement testing in kindergarten, first and/or second grades. The majority of the achievement test measures involved the Iowa Test of Basic Skills (ITBS) or California Achievement Test (CAT) with isolated studies using the Scholastic Achievement Test (SAT), Wide Range Achievement Test (WRAT), or locally designed measures. Teacher ratings, special education referrals, and grade retentions were also used as follow-up dependent variables study measures.

These 12 independent EPSF studies did not use the entire EPSF battery as a pre-test, usually eliminating the EPSF.MAS and the EPSF.PLS due to no published extensive standardization of these instruments until 1992. Even the intermittently released EPSF staff research on the program, typically used only the VMI, PPVT and EPSF.PLS for gain score comparisons in their preferred pre-post single group research design used through the mid 1980s. No independent studies were found that looked at using an independent variable involving the EPSF.PLS and EPSF.MAS subtests (total of 8) in predicting future reading success

Validity research has been done independently over the years on the PPVT and VMI assessment instruments. Previous independent research on the EPSF battery specifically involved only five studies in the literature with Terbush (1990), Bryant (1991), and Roth, et al. (1993) using canonical analysis or discriminant analysis to look at the predictive validity of the EPSF screening instrument. Roth, et al. (1993) was found to be the only independent EPSF predictive validity study that used the entire EPSF screening battery and EPSF modalities (but the researchers combined both Visual Discrimination (VD) and Visual Memory (VM) modalites into a single "Visual" modality for their research). McConnell (1986) did conduct discriminant analysis of the EPSF battery using 116 students involved in the EPSF treatment program. However, McConnell (1986) excluded the EPSF.DAP subtest results in the predictive analysis of the subjects' kindergarten year end EPSF and Metropolitan Achievement Test (MAT) testing.

The significance of the present study is to generate predictive validity research on the entire EPSF screening battery including all five tests, eight subtests and their resultant seven modality scores. The relationship of the EPSF screening battery to the Gates MacGinitie test, a widely used reading screening test, needs to be explored. Previously limited preschool kindergarten screening of predicting Gates MacGinitie has occurred.

The only documented attempt at EPSF screening battery comparisons to the Gates MacGinitie was done in preliminary research by EPSF staff during 1979 through 1982 (Strand & Werner, 1981). Computed Gates MacGinitie gain scores for defined "moderate" and "considerable need" kindergarten students were compared with a control group of non-EPSF treatment children. A one year gain in overall total Gates MacGinitie reading achievement test score was noted for the experimental versus the control group (Driscoll, 1992). However, the study was published in EPSF staff literature with limited discussion of the number of subjects in the study which was composed of "five pairs of experimental and control subjects from each school" (p. 37), but "in some cases less than five pairs were used" (Strand & Werner, 1981, p. 37). Other limitations of the Strand and Werner (1981) study included the use of the Gates MacGinitie Total Reading score without consideration of Gates MacGinitie Comprehension and Vocabulary scores and the exclusion of the EPSF.MAS, EPSF.PLS and EPSF.DAP subtests from the EPSF Battery, in effect eliminating 60% of the screening battery from the study.

Correspondence during 1993 with Dr. Werner and EPSF staff resulted in discovery of a ongoing current EPSF Project Office 1992-96 longitudinal study of the treatment effects of the EPSF intervention program with the Gates MacGinitie Reading Test as a major component of the study (see Appendix B). The present study by this author involves the predictive validity of the entire EPSF screening battery, its 8 subtests and 7 generated developmental modality scores. No current research on the treatment effects of the EPSF program was done in this study.

The current study would add to the anticipated developing literature on the relationship between the EPSF screening battery (its tests, subtests and resultant developmental modalities) and the Gates MacGinitie Reading Test. The three previous predictive validity studies of the EPSF Battery used: 1) limited samples with a maximum of 190 subjects, and 2) the most recently obtained EPSF screening data was in 1990 by Agostin (1993). Further current, more extensive predictive validity research on the EPSF screening tests and modality areas as compared to traditional reading achievement test results is needed due to the EPSF supplemental curriculum emphasis on reading skill development. More sophisticated predictive validity comparisons of EPSF screening data with future reading achievement is needed.

Problem Statement

The EPSF screening battery has been used for over twenty years as the primary defined "diagnostic" component of the overall EPSF public school intervention program designed to focus on reading and writing skill acquisition. Yet very limited research has been done on the predictive validity of the EPSF screening battery 5 tests ,its 8 subtests and its derived 7 modality scores as potential predictors of future reading success. Meisels, Wiske and Tivnan (1984) stated "most developmental screening instruments provide extremely limited validity information and very few describe the relationship between screening data and later school performance", (p. 25). No previous noted independent research on the entire EPSF screening battery, its subtests and modalities to predict any type of academic achievement could be found in the literature.

Previous noted independent research on the EPSF screening battery was overall critical of the EPSF generated EPSF.DAP, EPSF.MAS and EPSF. PLS lack of psychometric merit. No previous noted research was found that could effectively look at the potential predictive validity of these three EPSF generated screening battery tests. Previous research on the EPSF screening battery has frequently omitted the EPSF.MAS and EPSF.DAP data or minimized the contribution of these two tests in the EPSF research studies. Does the addition of the EPSF generated three screening test battery tests contribute to the predictive capabilities of the other two well documented EPSF screening tests included in the EPSF screening battery, namely the PPVT-R and VMI ?

No previous independent studies were noted in the literature regarding the predictive validity of the EPSF screening battery three MAS subtests and five PLS subtests. What is the extent to which the inclusion of the 3 EPSF.MAS subtests and 5 EPSF.PLS add, if any, to the predictive validity of the PLS Total test score and MAS Total test score for predicting future reading success? Do all or any of the 3 EPSF MAS or 5 EPSF.PLS subtests contribute significantly to the predictive validity of the EPSF screening battery?

Previous literature on the 7 EPSF derived modality scores has found varied significance of the predictive validity contributions of some of the modality scores. For example, Receptive Language (RL) has been discussed as a significant predictor of future reading success in some EPSF screening battery predicted reading tests studies. To what extent do all or any of the seven EPSF derived modality scores predict future reading achievement as measured by the Gates MacGinitie Reading Test?

Research Questions

The purpose of this study, conducted in a midwestern town, is to examine the predictive validity of the 1990, 1991 and 1992 school year entering Kindergarten students' results from the EPSF screening battery 5 tests, its 8 subtests and resultant 7 derived modality scores as predictors of future school achievement as measured by Gates MacGinitie (Form K, Level I) testing at the end of first grade. This study will examine:

1. What is the degree to which the kindergarten age administered EPSF screening battery 5 tests are related to and predict future, end of first grade, Gates MacGinitie reading achievement?

2. What is the extent to which the kindergarten age administered EPSF screening battery 8 subtests are relate to and contribute to the EPSF basic screening battery 5 tests' prediction of future, end of first grade, Gates MacGinitie reading achievement ?

3. What is the degree to which the kindergarten age administered EPSF screening battery generated 7 individual developmental modality scores are related to and predict future, end of first grade, Gates MacGinitie reading achievement?

Assumptions

This study is based on the assumption that developmental screening tests should have good reliability and predictive validity as part of their psychometric properties. Early detection of children with potential high risk for academic problems is assumed to be an essential prerequisite to maximize a child's academic success. Success in reading is assumed to be essential for overall academic success.

Anderson (1985) stated that the EPSF "developmental profile is used to determine the student's learning style and facilitate development of the classroom modality instructional program'" (p. 1). Effective prediction of potential reading success in kindergarten through diagnosis of a child's individual developmental style is assumed by EPSF test developers and some researchers in the literature to be valued educational information. It is assumed that the EPSF test developers base the majority of their EPSF program's component outcomes on the initial EPSF "diagnostic" program component that generates their developmental profile from the EPSF screening battery's derived 7 modality scores. Werner (1990) stated " the EPSF national validated program identifies each child's developmental levels and learning styles for the teacher initiated child centered, developmentally learning experiences in the classroom", (p. 1).

Limitations

This study deals only with a sample population of public school children from a midwestern community of approximately 35,000 people. Generalization of results beyond the identified sample population should be done with caution. The current study is limited to determining the statistical significance of the predictive validity of the EPSF screening battery results for determination of future achievement on the administered 1989 revised Gates MacGinitie Test

(Level 1, Form K) . Any generalization of these obtained results to other reading achievement measures should be done with caution.

The EPSF screening program recommends general hearing and vision screening, gathering obtained tester observations, and parental questionnaire information as part of the screening process. This information was not used specifically in the current study but is supposedly factored by the EPSF authors into the generated computer profile to determine a student's strength or modality need areas. Subsequent EPSF defined modality deficient skills are then recommended to be taught 15 to 20 minutes daily in the kindergarten classroom with EPSF designed curriculum materials . This study does not explore the effects of the EPSF kindergarten instructional program, only the EPSF standardized tests, subtests and resultant composite modality ratings in relationship to their prediction of future Gates MacGinitie Reading Test (Level 1, Form K) test scores.

This study involves 44 first semester kindergarten EPSF screening battery tested students' scores as well as EPSF screening results from pre-kindergarten entry children due to the fact that some students missed the summer EPSF screening and thus had to be EPSF tested later during the first semester of their kindergarten school year. Also, no documentation was available of the students in the current study potentially having had repeated kindergarten prior to their current study EPSF screening battery administration. Therefore potential generalized discussion of previous formal educational instruction background of all the subjects in this study can not be done from the current reported EPSF screening results due to (1) some of the EPSF screened children could have had previous educational exposure in retained kindergarten classes and (2) the fact that 44 current subjects were administered their EPSF screening during their first semester of kindergarten.

Organization of the Study

This dissertation is comprised of five chapters, references and appendixes. Chapter I includes an introduction of the rationale for early identification concerns regarding screening and screening instruments. The relevance of preschool prediction of future academic achievement is discussed. An overview of the EPSF screening program and the Gates MacGinitie Reading Tests is included. The significance of the study, problem statement, purpose of the study, research questions and organization of the study are discussed.

Chapter II is a survey of literature involving kindergarten screening, an overview of kindergarten screening tests, predictive validity in kindergarten screening tests, kindergarten screening test prediction of later reading achievement (including a discussion of the Gates MacGinitie Reading Test), the EPSF screening battery and its intervention program as well as a summary of the major points noted.

Chapter III is the discussion of the research methodology. It includes a brief outline of chapter content followed by discussion of study subjects, instrumentation, data collection, procedures and statistical analysis preformed.

Chapter IV is the statistical analysis of the study data. Chapter V contains a summary of the study purpose, methodology, research findings, conclusions that can be drawn with recommendations for further research discussed. Chapter V is followed by references and appendixes.

Chapter II

Review of the Literature

Overview

This chapter reviews relevant literature on kindergarten screening, limitations of kindergarten screening, an overview of kindergarten screening tests, predictive validity in kindergarten screening tests, kindergarten screening test prediction of later reading achievement (including the Gates MacGinitie Reading Test) and the EPSF screening battery and program (with emphasis on EPSF staff research summative reports and independent research on the EPSF screening battery, its subtests and generated EPSF modalities).

Kindergarten Screening

The use of kindergarten screening has increased dramatically in the last twenty five years as research on preschool development of readiness skills has evolved. Federal and state legislation began in the mid 1960's increased the awareness of the need for early identification and intervention with younger children. Rapid growth of preschool educational attendance has occurred in the last twenty five years from about 15% of four years olds in 1967 to approximately 50% of all U.S. four years olds in 1986 (Slavin, Karweit & Madden, 1989). Thus, the growth in preschool education can be attributed, in part, to increased public awareness and research in the literature resulting in increased federal and state legislation. Significant federal mandates included the Elementary and Secondary Education Act of 1965 creating Headstart, legislation creating EPSDT in 1967, PL 93-80 creating Child Find in 1974, PL

94-142 in 1975, and PL 94-457 in 1986 (Paget, 1990).

"Kindergarten attendance is almost universal in the United States today" (Slavin, et al. 1994, p.79). More than three million children enter kindergarten every year (Shephard & Smith, 1986). The increased awareness and positive influence of early intervention has helped increase kindergarten attendance in the United States to 93% of five year olds who are enrolled in school today (Slavin, et al. 1994). Mandatory kindergarten attendance is required in twelve states and the District of Columbia (Slavin, et al. 1994).

This increased emphasis of formal school attendance for kindergarten has increased the need for effective kindergarten screening programs. These screening efforts have been done with the best intentions of the majority of preschool screening program and screening test developers. The emphasis on early intervention for young children is based on the assumptions that early intervention is better for changing children, can potentially decrease the magnitude of developmental problems, and possibly reduce cost factors through early detection of potential difficulties. In effect, Harrison (1993) stated "prevention is more effective and more economical as a rule than repair, better to identify problems early and correct them promptly than to let them grow into crisis requiring action".

Screening efforts for preschool and kindergarten children have increased dramatically in the last two decades. "More than 25 states currently mandate developmental screening for three to six year olds" (Meisels ,1987, p.6). Minnesota in 1977 was the first state to require comprehensive free screening for all kindergarten age children (currently 45,000 screened each year), as noted by Nuttall, Romero and Kalesnik (1992). Gracey, Azzara and Reinherz (1984) stated that 22 of all 50 states have required preschool or kindergarten screenings and Nuttall, et al. (1992) reported 16 states have required kindergarten screenings. This heightened emphasis in the last two decades on preschool screening has evolved into more widespread kindergarten screening and school readiness testing.

The increased demand for the creation of (1) preschool and kindergarten programs and (2) kindergarten screening instrumentation has lead to statements in the literature of best practices in kindergarten and preschool screening. McConnell (1986) stated that kindergarten screening should make effective use of the concept of developmental age, be widely accessible, systematic, quick and simple with an aggressive child find component. Miesels (1985) stated that screenings should lead to adaption of the system, not the child, with emphasis on identifying individual traits in the child related to later learning. Therefore, Miesels (1985) emphasized that screening tests should be not be used for school entry but to identify traits related to later learning.

Specific benefits of screening preschool children noted by Bailey and Wolery (1989) include individualized program planning, diagnostic placement, program placement, evaluation decisions, and differentiation of at risk as compared to normal preschool children. Graue (1993) stated that the National Governors' Association in 1990 established a set of recommendations for U.S. education headed by the objective that "by the year 2000 all children in America will start school ready to learn". Despite the notable needs for and benefits of preschool screening, there exists "no acceptable standards that allow for universal comparison of children's progress", (Mcloughlin & Rausch, 1990). Preschool test and program developers have generated varied materials to meet the demands for effective early identification and intervention with preschool children.

Limitations of Kindergarten Screening

The initial enthusiastic efforts by schools and other social agencies toward early identification and remediation of potential learning problems has resulted in some definite criticisms of screening limitations over the years in the research literature. Mcloughlin and Rausch (1990) stated "It is likely most, if not all, the data obtained from childrens' screenings are underestimates" because of unsophisticated test taking skills by preschoolers, typically heavy loading of screening verbal items with some younger children having undeveloped verbal skills, test subject motivational errors and higher incidence of test rapport issues with preschoolers.

Specific screening limitations of children mentioned in the literature include: (1) premature labeling (Adelman,1982; Barnes, 1982; Garner, 1993; Lichenstein & Ireton, 1984; Mercer, Algozzine & Trifiletti, 1979; Paget & Nagle, 1986; Thurlow, et al. 1986) (2) limited generalization of results (Lichenstein & Ireton, 1984; McGowen, 1989; Miller & Sprong, 1986); (3) difficulty handling the issue of rapid developmental change in preschool children; and (4) psychometric difficulties in preschool screening tests (Adelman, 1982; Barnes, 1982; Bailey & Wolery, 1989; Lichenstein & Ireton, 1984; Meisels, 1985 & 1989; Rosenkoetter & Wanska, 1992; Satz & Fletcher, 1988). Adelman (1982) commented on widespread application of screening procedures as "another example when pressure and enthusiasm for new procedures have led to inappropriate extrapolation of research findings and premature applications", (p. 255).

The issue of labeling preschool children as a result of the screening process has been widely discussed in the literature. McConnell (1986) comments on labeling misuse in kindergarten screening as being inappropriate due to the concept of self fulfilling prophecy which places overemphasis on describing a child not yet exposed to formal education. McConnell (1986) stated Koegh and Becker (1973) made the comment that screening test results are "hypothesis about future development based on present performance" (p. 16). Barnes (1982) stated that "screening measures are not designed to be that precise or specific, rather they are designed to be administered singly or together in a battery with the single objective of detecting children at-risk" (p. 34). Garner (1993) adds that kindergarten screening programs are "only intended to identify children at risk of experiencing academic difficulties or those children who may benefit from instructional assistance" (p. 128) and labeling a student can reduce their opportunity to participate in a regular educational setting.

Bryant (1991) stated the issue of rapid developmental change in preschool children is noted in Miller and Schouten's (1988) concept that difficulties developing valid screening tests are compounded by the unpredictability of early child development. Barnes (1982) noted that children frequently outgrow screening identified deficits without special education intervention. Bryant (1991) states that Barnes (1982) has noted in preschool children's rate of development "large inter- and intra-individual differences in rate of growth in various developmental domains" (Bryant , 1991, p.11) . Werner (1990), developer of the Early Prevention of School Failure Project (EPSF) screening battery and programs, stated from over 50,000 EPSF screening profiles gathered from 1974 through 1989, that a bimodal frequency distribution exists of preschool children's abilities who are coming to kindergarten. Werner (1990) noted that children enter kindergarten with 40% having "advanced skills", 40% with developmental delays and the 20% of "average" children "disappearing in many schools (p. 2).

Other limitations of preschool screening include difficulties in detecting children with mild developmental delays. Mercer, et al. (1979) see kindergarten screening as a gross measure of functioning more effective in looking at extremes of functioning. This is due, in part, to the lack of representative sampling in most screening tests standardization samples of non-normal children. Therefore, it is difficult to interpret developmental rates of handicapped children from screening test data.

Kindergarten Screening Tests

The enthusiastic efforts to identify at risk children has resulted in some theoretical as well as practical issues regarding screening tests. Some question has been raised in the literature over differences between screening testing, diagnostic testing and readiness testing. Screening tests are not for diagnostic purposes. Adelman (1982) stated that it is "not uncommon for screening instruments to be misused. Some screening instruments generate labeling of children which labels, in turn, can be interpreted as diagnosis for prescribed intervention" (Adelman, 1982, p. 258). Satz and Fletcher (1988) cautions that screening "should not be confused with diagnosis" (p. 825) and should be used as a quick, cost effective system not requiring professionals interpretation for children at risk for subsequent difficulties. Lichenstein and Ireton (1984) comment:

> Many screening instruments are simply brief versions of comprehensive assessment measures developed for diagnostic not screening purposes. Thus, circumventing complications simply by changing the number of test items and reducing administration time (p. 123).

Adelman (1982) states first level screening is intended to survey large groups in the first stage identification process and to detect problems rather than designate procedures for diagnostic classification. Meisels (1985) in his excellent discourse on screening versus diagnostic assessment clarified screening as a limited procedure to "select children who may have special needs and not to label, pace or develop intervention procedures" (p. 5). Meisels (1985) conceptualizes diagnosis as a process to identify children who have special needs with focus on the nature of the problem with suggested causes and appropriate remedial recommendations.

Confusion of kindergarten screening versus readiness testing has been noted by Gridley, Mucha and Hatfield (1995) and Meisels in articles dated 1985, 1987 and 1989. Meisels states that "even the Burros Mental Measurements Yearbook does not distinguish between readiness and screening" (Meisels, Wiske & Tivnan, 1984, p. 1). "Substituting readiness for screening testing occurs inadvertently through confusion over the difference between them (Meisels,1987, p. 6).

Meisels (1987) main distinction between screening versus readiness involves the fact that screening test content looks at " a child's ability or potential to acquire skills" (p. 5) with the purposes of identifying children who"may need early intervention or special education services" or "might profit from a modified or individualized classroom program" (p. 5). He conceptualizes readiness testing content as focusing on "current skill achievement, performance and general knowledge" with the purposes of facilitating curriculum planning and identifying "a child's relative preparedness to benefit from a specific academic program" (Meisels, 1987, p. 5).

Meisels (1987) perceives predictive validity as a major issue for screening tests due to their focus on learning potential versus reading readiness requires more focus on construct validity due to looking at the child's current achievement or performance. Meisels (1987) summarizes his position on readiness versus screening by stating that "Fixing readiness problems leads policymakers to increased frequency of adopting screening programs for at risk children which leads to screening focusing on readiness/ developmental immaturity" (p. 5).

Confusion of readiness testing concepts used in kindergarten screening testing can lead to premature prediction or labeling of children's learning potential. Agostin (1993) states that at times developmental age in readiness tests is used to determine kindergarten or first grade readiness. Charleswood (1989) speaking on the negative effects of kindergarten screening stated "Often what happens is rather than provide the child with optimal experiences needed, the child often ends up further behind" (Agostin, 1993, p.4). Repeating Grades (1990) research mentioned by Agostin (1993) commended that approximately 5 to 7 percent of United States public school children are retained each school year with as many as 50% of kindergarten children students retained in some school districts. There is "no evidence that kindergarten retention, developmental kindergarten or transitional first programs are more effective then simply promoting children" (Slavin,et al. 1994, p. 119). Thus, screening tests should not be used to predict future grade placement nor premature prediction of learning potential.

Concerns of the technical merits of preschool kindergarten screening instruments are readily evident in the literature (Barnes, 1982; Bracken, 1987; Bryant, 1991; Lichenstein & Ireton, 1984; Lindsay & Wedell, 1982; McConnell, 1986; Meisels, 1987; Reynolds & Kamphaus, 1990). The American Psychological Association (APA) even as early as 1974 established dimensions on which screening tests should be evaluated including specified: (1) normative sample; (2) sample size of 100 subjects for each subgroup; (3) systematic item analysis; (4) reported measures of central tendency; (5) test retest and interrater reliabilities of .9 , with (6) statistical significance beyond .05 reported for concurrent and predictive validities and (7) test manual reported test procedures and examiner qualifications (Bailey & Wolery, 1989). Other researchers state specific requirements for effective screening measures should include test-retest reliabilities of .8 (if used for individual decisions) and inter-rater reliabilities of .8 (Bracken, 1987; Lichenstein & Ireton, 1984; Lehr, Yysseldkye & Thurlow ,1987; Rosenkoetter & Wanska,1992; Salvia & Yysseldkye, 1991).

Bracken, (1987) stated since the advent of P.L. 94-142 and P.L. 99-457, the significant increase in testing of preschool children and number of preschool assessments has led to the need for "increased professional attention paid to the quality of instruments used in preschool assessment" (p. 314). Complicating this issue is the fact that Meisels (1985) states that the 1985 APA generated guide for preschool and educational testing indicates "screening tests should only be used if they meet acceptable criteria of standardization relationships and values" (p. 3). Bracken (1987) stated "these standards are, in many cases, too general and do not set criteria for special areas of technical adequacy" (p. 313).

Specific issues of technical adequacy for kindergarten screening tests include subtest item gradients, subtest and total test internal consistency with special emphasis regarding test floors and validity concerns (Bracken,1987). Satz and Fletcher (1988) state sample size should be large enough to handle subject attrition and be representative of the sample population yet unfortunately "most screening instruments are standardized on samples of 30 to 60 subjects with typical follow-up intervals of one year or less" (p. 826). One hundred subjects or more in any sample per age or grade is recommended by Lehr, Yysseldyke and Thurlow (1987).

Concern regarding subtest and total test floors by Bracken (1987) focused on his position that minimal levels should be established for differentiation of low functioning children and low to low average children. Reynolds and Kamphaus (1990) voice concern regarding screening instrument ceilings due to less stability in higher scores on the upper end of the screening scale due to the fact that upper end items are worth more than corresponding items at the bottom of the scale.

Bracken (1987) states that preschool instrument generated subtest item gradients are not effective due to large changes in children's obtained screening results caused by a single score. Preschool screening measures typically have large standard score differences in relationship to changes in raw scores thus cause the instrument to be less sensitive to small changes in preschool children's abilities (Bracken, 1987). Some validity research has found evidence that preschool tests have greater predictive accuracy or defined higher correlations for predicting low functioning children (Lindsay & Wedell, 1982; McConnell, 1986; Paget, 1990; Roth, et al. 1993).

Previous extensive discussion of the inherent risks or limitations of preschool screening results utility has been done (Adelman ,1982; Barnes, 1982; Harrington, 1984; Lindsay & Wedell, 1982; Meisels, 1985; Satz & Fletcher, 1988). For example, Satz and Fletcher (1988) state one of the most frequent problems with preschool and kindergarten screenings is "inadequate assessment of predictive utility of screening device" (p. 24).

The two key error patterns in preschool and kindergarten screenings utility involve identifying a child at-risk when no problem exists (false positive) or failing to identify a child who has a potential problem (false negative). Meisels (1985) stated that most validity studies of screening instruments involve correlational analysis describing the degree of overlap between two tests or measures, thus, yielding no information about the accuracy of the screening test results or the number of children over referred or under referred. Harrison (1993), Lichenstein and Ireton (1984) and Meisels (1985) describe the relationship of false positives to false negatives as the "hit rate" of the preschool screening instrument. They state the lower the percentage of false positive and false negatives, the more accurate the screening test or procedure. Thus, screening test developers face the dilemma of developing quick, cost effective instruments for general screening of typically larger groups of individuals yet still maintaining the technical adequacy and test utility necessary for predicting preschool and kindergarten children's abilities and needs.

Thus, statistical theory for construction of a good screening instrument has been known and discussed. Still the technical merits, especially regarding the reliability and validity issues, abound in the literature regarding the rapidly generated number of preschool and kindergarten screening tests credited over the years to potentially assess young children's learning abilities. Meisels (1989) stated, "developmental screening tests are in widespread use but few reliable and valid tests are available" (p. 578).

Joiner (1977) conducted a survey of 177 New York State school districts and found 151 different procedures and or tests utilized for preschool screening with only 16 having "even marginal reliability and validity". A 1971 UCLA published comprehensive evaluation guide of over 120 preschool and kindergarten tests (having 360 total subtests) found only seven subtests "rated as providing good measurement validity" (Kelley & Surbeck ,1983, p. 12). Lehr, Ysseldyke and Thurlow (1987) evaluated 109 different preschool tests used by 54 United States Early Education Programs for Handicapped Children (ages birth to six years). They found that of the 19 most used tests only three "had technically adequate norms, validity and reliability" (Lehr, et al. 1987, p. 397).

Meisels (1987) reported a 1984 Michigan Department of Education survey of 111 tests used for preschool, kindergarten and pre-first level programs. He stated "fewer than ten of these tests were appropriate in terms of age and purpose to which they were put to" (Meisels, 1987, p. 5). Nuttall, Romero and Kalesnik (1992) reported that of eight reviewed screening tests only four were found to be standardized nationally and appropriate for most children.

Adelman (1982) commented that:

> Screening is much in demand especially for identifying learning problems at an early age. A climate has been established when both consumer and suppliers are less critical than they should be in evaluating the validity of proposed and previous procedures. (p. 25)

Validity of preschool and kindergarten screening instruments has been discussed indepth by Barnes (1982) and Stangler, Huber and Routh (1980). Bracken (1987) in his discussion of preschool instruments technical adequacy focused his dialogue on reliability versus validity due to the fact that "any given test can only have a validity coefficient that is as high as the square root of the reliability" (p.325). Thus, if either the predictor instrument or the criterion instrument have low reliability, then the correlation between the two instruments will be lower than if the instruments were both reliable. Nevertheless, validity of kindergarten instruments is an relevant issue. Meisels (1985) stated that some screening tests report results in terms of face validity by use of "independent judgment of professionals concerning the relevance of a screening instrument" (p. 12). He states this is an imprecise method that does not imply administration of further empirical research and "should not be used as a substitute for other validity procedures" (Meisels, 1985, p. 12).

Lehr, Yysseldkye and Thurlow (1987) stated a common evaluation criticism of screening test technical adequacy involves the lack of validation reported in the test manual or accompanying technical publication including discussion of at least one type of validity, be it content, construct, or criterion related. Meisels (1985) stated that concurrent and predictive validity were the two types of validity most reported in screening instruments. Bryant (1991) states that screening tests are developed to be quick and cost effective but "have inherent risks due to not having indepth or extensive validation procedures" (p. 11).

Predictive Validity of Kindergarten Screening Tests

Bailey and Wolery (1989) refer to predictive validity of screening tests as " the extent to which the screening test agrees with the child's performance or outcome measures later in time" (p.127). Satz and Fletcher (1979) reported the importance of predictive validity issues in preschool and kindergarten programs. He stated "the importance of predictive value of preschool screening programs is directly related to the predictive error rate of the measures employed" (Satz & Fletcher, 1979, p. 45).

Lindsay and Wedell (1982) in their discussion of screening instruments and their predictive validity capabilities stated:

> While it is to be expected of instruments seen to be new and experimental that a small amount of information is available to evaluate them, it is worrying when instruments are used up to ten years with still very little evidence of of their usefulness. (p.214)

Lichenstein and Ireton (1984) stated the value of a screening instrument includes psychometric qualities, especially the predictive validity issue. McConnell (1986) in her study of the predictive validity of the EPSF screening test battery stated in her review of the literature that "Most predictive studies utilize correlation techniques to determine relationships between screening test performance and achievement tests are administered at the end of the school year. Moderately high correlations of .50 to .80 have frequently been found." (McConnell, 1986, p. 23).

The incidence of studies in the literature regarding preschool and kindergarten screening tests predictive validity limitations is well documented (Joiner, 1977; Lehr, Ysseldyke & Thurlow, 1987; Meisels, 1987). More recent predictive validity studies of preschool and kindergarten screening instruments include Ellwein, Walsh, Eads and Miller (1991) study of four preschool screening instruments. Ellwein, et al. (1991) found all four tests to have lower predictive validity (Graue, 1993). Thus, predictive validity continues to be a current concern for preschool and kindergarten screening instruments

Some researchers have critized predictive instruments due to the lack of available reported validity information (McConnell, 1986). Lehr, Ysseldyke and Thurlow (1987) commented that the Standards for Educational and Psychological Tests (APA, 1985) stated "validity should be reported in the manual or in an accompanying technical manual" (p. 395). It is also stated by Lehr, et al.(1987) that the 1985 APA criterion for technical adequacy of tests clearly stated for predictive validity "a statement concerning length of time for which predictions can be made should be included" (p. 395). The seriousness of predictive validity concerns is clearly noted by Meisels' comment that use of screening tests without validity data is "an abuse of testing procedure and of the trust the community places in professional educators" (Miesels ,1987, p. 6).

In contrast, Miesels (1985) stated that even kindergarten and preschool screening tests with good validity show "marked decline in accuracy of prediction over a two year or more period " (p.29). The predictive validity limitations of preschool and kindergarten screening tests are apply summarized by Meisels (1985) comments that:

> With the criteria a developmental screening instrument must satisfy- brevity, efficiency, low cost, standardized administrating,

> objective scoring, non- diagnostic focus, development content, validity measured by classification rather than correlational methods, the possibility of long term predictive accuracy may be unattainable. (p. 29)

Despite the potential limited predictive accuracy of preschool and kindergarten screening results beyond a period of two years, the fact remains that this two year period for a kindergarten student is one of the most critical times in a child's academic life. Recent research in academic expectations for early elementary school children has shown that increased expectations for academic skills at earlier grade levels is occurring (Agostin, 1993 and Slavin, et al. 1994). Charlesworth (1989) noted that kindergarten in the 1970's emphasized learning through play and socialization for developing school readiness skills. Agostin (1993) commented that Charlesworth (1989) found in 1980's the trend toward increased preparation in kindergarten to meet first grade curriculum demands through increased academic curriculum in kindergarten.

Kindergarten Prediction of Later Academic Achievement

Kindergarten screening of academic skills, especially reading, has increased dramatically in the last two decades as increased stress in formal academic learning for kindergarten children has occurred (Charlesworth, 1989; Shephard & Smith, 1988; Slavin, et al. 1994). Slavin, Karweit and Madden (1989) stated "Most kindergarten programs in public schools are focusing either directly on academics (22%) or on academic preparation (63%)" (p. 103).

Shephard and Smith (1988) reported a 1986 Educational Research Services survey conducted with school administrators and teachers. They found formal reading instruction in kindergarten classrooms noted by 18% of school principals reporting it was school district policy to teach reading to all kindergarten children with approximately 50% of reporting schools teaching kindergarten children who were "ready and able" to read (Shephard & Smith, 1988). Increased stress on kindergarten formal academic curriculum has lead to the increased use of developmental first and transitional first grade programs in the United States (Shephard & Smith, 1988 and Slaven, et al. 1994).

Previous research on preschool and kindergarten screening measures as predictive measures of subsequent academic success includes a "proliferation of studies which relate either a single screening test or a battery of tests to subsequent achievement" (McConnell, 1986, p. 23). McConnell (1986) stated previous researchers as early as Evans and Ferguson (1974) have placed most predictive measures into one of three categories including reading readiness tests, measures of general academic or school readiness and identification of learning disabilities or learning potential. Meisels, Wiske and Tivnan (1984) stated that "most developmental screening instruments provide extremely limited validity information and very few describe the relationship between screening data and later school achievement", (p. 25).

Kindergarten screening testing to predict later reading success has focused on defining the predictive variables noted in screening measures proported to measure prerequisite skills necessary for later achievement. Screening instruments are similar in their content usually having five to six subtests focusing on different aspects of child development, including language, visual and auditory perception, motor skills, perceptomotor functioning and letter recognition (Lichenstein & Ireton, 1984).

Barnes (1982) stated the "basic objective of screening is to identify as soon as possible those preschoolers who, for whatever reason, do not seen to be adequately developing those skills necessary for later academic screening" (p. 175). Barnes, (1982) states the critical skill areas needed for eventual reading skill development include auditory, visual, auditory-visual language, rate of learning words, the concept of reading and reading rate.

Some of the initial research on predicting later reading success was done in the 1930's by Durrell and Sullivan through their research on language as a predictor of later reading success. Learning rate of words since the 1950's has "consistently been a fairly powerful predictor of later reading achievement" (Barnes, 1982, p. 177). He states research in the 1970's looked at specific auditory and visual skills including auditory discrimination, auditory blending, rhyming , auditory memory was well as visual discrimination and visual memory as relevant screening areas for predicting future reading success (Barnes, 1982).

Some of the previous notable research in individual kindergarten screening test or test batteries includes deHirsch, Jansky and Langford (1966) study of 37 different tests and their correlation with later reading achievement. They determined that only two tests, the Metropolitian Reading Test and the Bender Gestalt received a correlation of at least .5 with future reading success. "Knowledge of letter names was determined (by deHirsch, et al. 1966) to be the best single predictors of reading achievement" (McConnell, 1986, p. 24).

Mercer, Algozzine and Trifiletti (1979) did a survey of 15 studies from 1970 through 1977 involving prediction of kindergarten and first grade children's future academic success from measures gathered eight months to as long as seven years later. The kindergarten studies discussed by Mercer, et al. (1979) involved a range of 26 to 572 subjects. The predictive utility of obtained predictive variables for future academic achievement yielded median hit rates ranged from 75% for single measures to 79% for test batteries and almost 80% for teacher ratings. Limitations of the Mercer, et al. (1979) study included no description of the 15 studies given by the authors.

Horn and Packard (1985) conducted a meta-analysis of 58 studies dating 1960 to 1980 regarding prediction of reading achievement. "Correlation between measures administered in kindergarten or first grade and reading achievement later in first to third grades in elementary school " was conducted (Horn & Packard, 1985, p. 597). The various 58 study predictor variables noted by Horn and Packard (1985) included: (1) language areas including written, oral expression and receptive; (2) sensory area including figure drawing, auditory and visual perception; (3) sensory integration; (4) behaviorial-emotional area including attention distractibility, externalizing, internalizing, self help and social skills; (5) soft neurological variables including fine/gross motor and cerebral dominance/handedness; (6) an IQ measure and 7) teacher ratings.

Horn and Packard (1985) in their analysis found the best overall kindergarten or first grade predictor of later reading achievement involved the two behavioral-emotional variables of attention/distractibility (mean r of .63) and internalizing (mean r of .59). The next most highly rated overall predictor variables included written expression (mean r of .58), receptive language (mean r of.56) and group IQ tests (mean r of .55). The best predictors by variable areas were IQ (mean r of .53) and language (mean r of .52) with sensory, teacher ratings, behavioral-emotional, and soft neurological signs receiving mean r of .42, .49, .48 and .41, respectively. It is interesting to note that teacher ratings ranked as a better predictor than any motor or sensory (e. g., auditory and visual processing skills) predictor variables.

Tramontana, Hooper and Selzer (1988) conducted a meta-analysis of 74 studies dating from 1973 to 1986 involving only kindergarten children predictor variables as measures of academic success. A time interval of at least one year was required between obtained initial kindergarten predictor variables and follow-up measures for any study to be included in the meta-analysis. Tramontana, et al. (1988) generated a vast array of information on kindergarten children variables including, IQ/general cognitive abilities, specific cognitive abilities, language skills, perceptual/perceptual motor skills, behavioral - emotional functioning and demographic factors as predictors of later behavioral-emotional and academic achievement in reading and math areas.

Tramontana, et al. (1988) found significant kindergarten predictor variables of later reading success were "cognitive, verbal and perceptual/perceptual motor measures and prediction probably strengthened when measures from each of these categories are combined" (Tramontana, et al. 1988, p.131). They found different predictor variable patterns for different grade levels due to different developmental factors potentially required at each grade and curriculum level. Cognitive and verbal kindergarten predictor variables had lower predictive power for reading achievement until second and third grades "possibly because reading at a beginning level depends more on perceptual recognition abilities" (p. 132). The defined language predictive variable in many studies was found to be the best predictor of first grade reading.

Tramontana, et al. (1988) found the best single measure predictors of first through third grade reading were (in descending rank order) letter naming, general cognitive ability, language, visual motor and finger localization. Thus, they found letter naming/reciting was the "best predictor for later reading achievement' (p. 127). Tramontana, et al. (1988) found language (both receptive and expressive) in multi-measure assessments often was among the best predictor of reading and math achievement. "Visual-perceptual and visual-motor measures contribute effectively to the prediction of reading, math and general achievement at least through first grade" (Tramontana, et al., 1988 , p. 127). Fine and gross motor skills were noted only in a few of the 74 studies as predictor variables and were not seen as good kindergarten predictors of future reading success. Verbal abstraction was noted as a specific cognitive area effective in prediction of later reading success.

Tramontana, et al. (1988) found increased academic prediction for second grade and higher due to lower stability of academic skills until the end of first grade. Tramontana, et al. (1988) referred to the Butler, Marsh, Sheppard and Sheppard (1985) study stating that 'Whereas measures of preschool abilities directly predict a child's initial success in reading, it is the child's actual achievement in the first or second grade that is directly predictive of achievement in later grades" (Tramontana, et al., 1988, p. 134).

Tramontana, et al. (1988) overall summarized by stating: (1) with exception of children with significant noted disabilities, the optimum time for initial screening of the general population of preschoolers children "would be roughly at the end of kindergarten" (p. 139); (2) hit rate concerns, especially for false negatives with a caution some children functioning the middle range " can grow into a deficit over time" (p. 138) and (3) multi-assessment kindergarten batteries versus single measures should be used as better predictors of future academic success (Tramontana, et al. 1988).

Early predictive studies of later academic achievement have looked at individual highly correlated variables frequently developed into screening batteries, such as the EPSF discussed in this study. Horn and Packard (1985) state that:

> Much of the empirical literature has been concerned with identifying early predictors of later school success and failure. In general, these studies have correlated motoric, cognitive, perceptual, sensory, and behavioral variables assessed in kindergarten or first grade with later school achievement.The variables with the largest correlations with future school achievement were then defined as providing the best early prediction of future academic status. Subsequently, the later variables are often incorporated into early screening batteries for the identification and eventual treatment for children at high risk for the development of learning problems in school. (p. 597)

Wallbrown, Engin, Wallbrown and Blaha (1975) study, using a multi-instrument screening battery, was one of the few predictive validity studies in the literature that used a multifaceted view of reading achievement versus the frequently used total composite reading test score. Wallbrown, et al. (1975) study involved preschool prediction of first grade Gates MacGinitie reading achievement as measured by the Gates MacGinitie Reading Tests Vocabulary, and Comprehension scores.

Wallbrown, et al. (1975) found that the best single predictor of first grade

Gates MacGinitie Vocabulary and Comprehension subtest were the Slosson IQ Test (accounting for 28% of the variance) and the Bender Gestalt (accounting for 38% of the variance), respectively. The Gates MacGinitie Vocabulary subtest overall was better predicted than the Gates MacGinitie Comprehension subtest. Wallbrown, et al. (1975) commented that "both IQ and visual motor integration are important components of first grade reading" (p.148). They stated that good first grade reading comprehension requires a broader range of visual skills than reading vocabulary. Study limitations include only a total of 100 subjects in study.

Some predictive validity research on kindergarten and preschool measures or instrument finds that in reading achievement different processing skills are relevant as predictors of future reading achievement at different grade levels. Barnes (1982) stated that in short term prediction of reading disability, an outcome measure "may be tapping different constructs at different grade levels" (p. 30) with noted increased difficulty of ceiling level test items at higher grade levels. Greenfield and Scott (1985) stated researchers need to look at subskills of different domains.

Thus, further research is needed on specific reading achievement areas at different grade levels and the relationship of specific different domains subskills as noted on early screening instruments or batteries. Still, Lindsay and Wedell (1982) caution as diagnostic focus shifts down the age range the type of process or ability investigated becomes more remote from the target task, e.g. reading. This caution also holds true with potential shifts in children's age of screening processing skills as measures of future academic success.

The literature shows that further research is needed on kindergarten screening measures and the theoretical assumptions underlying the constructs. Slavin et al. (1994) stated the increase in kindergarten preparation for school and the role of kindergarten needs to be explored. Graue (1993) states the need for readying children for kindergarten should include increased emphasis made in skill deficit models with skill building, not test taking, a priority.

Numerous preschool and kindergarten screening programs with corresponding teaching components have been developed to get students off to a good start (Slavin, et al. 1994). One of these programs is Early Prevention of School Failure (EPSF) in existence since 1971. The EPSF program uses a kindergarten screening battery as the basis of its diagnostic component to generate both kindergarten and first grade intervention programs to help prevent reading failure (Werner, 1990).

Early Prevention of School Failure

The Early Prevention of School Failure (EPSF) is a nationally validated diffusion program designed to prevent school failure through early identification of four to six year old children's developmental skills and learning style (Werner,1990). Werner stated that EPSF "began in 1971 in southern Wills County, Illinois as a Title III ESA Project" in response to 1969 legislation requiring special education services for children ages 3 to 21" (Werner, 1990, p. v).

The EPSF project was nationally validated in 1974 by the United States Department of Education funded Nation Diffusion Network (NDN) and Joint Dissemination Review Panel (JDRP) as an exemplary identification and developmental program for children four to six years old. Follow-up NDN validation of approved programs is required at least every six years to continue in an exemplary program status. The EPSF program in 1977 was JDRP validated for use with Chapter 1 and migrant children. NDN/JDRP program validation occurred again in 1985 and in 1990. The EPSF first grade curriculum entitled "Success in Reading and Writing" was initially NDN validated in 1990.

The JDRP or (as it was known after 1987) the Program Effectiveness Panel (PEP) is the program evaluation component of the NDN (Educational Programs That Work,1994). The NDN catalogue description of approved programs known as Educational Programs That Work (EPTW) still lists in 1994 the EPSF as an exemplary program. NDN exemplary status requires a program to have an objective evaluation of its effectiveness" submitted by the developer of the program" with the criterion for JDRP/PEP panel members that they should be "convinced that the program has meet its stated objectives at the original development or demonstration site" (EPTW, 1994, p. 9).

Slavin, et al. (1994) lists the EPSF as only one of seven original NDN approved exemplary kindergarten projects still active. Slavin, Karweit and Madden (1989) cautioned that program developers typically only give limited description when trying to meet JDRP/PEP mandates with no strong evaluation design component required by JDRP/PEP (such as random assignment of subjects and treatment/control groups). Slavin, et al. (1989) states many of the JDRP/PEP studies "should be viewed as illustrations of possible effective strategies and should be candidates for a more through evaluation" (p. 89).

The EPSF program has been reported to have "over the last twenty years been piloted in over 2,000 school districts located in 48 states and in five foreign countries " (Driscoll ,1992, p. 18) with "439 certified trainers in 49 states, the Virgin Islands and Canada" (Werner,1990, p. vi). Werner in 1987 stated the EPSF "presently serves over 500,000 young children" (Bryant, 1991). The EPSF program reportedly has received federal (including Title II, Title IV and NDN) and state funds (state of Ohio in 1976 and Hawaii in 1982). Thus, the EPSF program has received widespread use and support.

Program Components

Werner (1990) stated the EPSF screening battery was just one of six component parts of the overall EPSF program. The generation of the EPSF developmental profile is essential for the instigation of the followup EPSF program components. The EPSF screening battery is used to generate the EPSF developmental profile. Further critical review of the EPSF screening battery is needed to justify its role as a major tenet of the overall EPSF program. Werner (1990) defines the six components of the EPSF program as:

1) "Diagnosis" looks at "the child's developmental levels and preferred learning style" (Werner ,1990, p. 8). The EPSF Diagnosis component includes team screening using the EPSF screening battery with follow-up team conferencing of the obtained EPSF computer generated child's "diagnostic student profile" (see Appendix A) which delineates the child's relative strengths or needs and potential for being at risk in seven developmental areas.

2) "Curriculum design" is based on "observation and screening information" (p. 8) which is noted in the EPSF 52 identified "critical and observable developmental skills that provide the foundation for reading and math skills" (Werner, 1990, p. 8). These noted EPSF objectives generate learning activities for use by the classroom teacher in direct EPSF modality instruction up to 15 to 20 minutes daily for children identifying as at- risk of learning failure in one of the defined EPSF seven developmental modality areas.

3) "Classroom management" involves a EPSF generated format for systematic record keeping on EPSF developmental modality instruction and the individual student's progress in their individual specified developmental modality need areas.

4) "Parent Involvement" involves encouragement by the classroom teacher of parents to "become knowledgeable about the program, to volunteer in the classroom and to work with the child at home" (Werner, 1990, p. 8).

5) "Evaluation" involves the EPSF staff program developers stated philosophy of continuing "to evaluate the effectiveness of their training workshops, follow-up inservices and total replication of the program" (Werner, 1990, p. 9). Supposedly "educational agencies can participate in a project sponsored three year longitudinal study, annual evaluation study or initiate a local research study", (Werner,1990, p. 9).

6)"Inservice Training" involving basic and advanced levels for project implementation staff and EPSF parents. An excellent overview of the EPSF components is found in Driscoll (1992).

The EPSF screening battery is compromised of five standardized tests that generate seven developmental modalities (Werner, 1990). The five EPSF screening battery tests include the Peabody Picture Vocabulary Test- Revised (PPVT-R), Developmental Test of Visual Motor Integration (VMI) as well as the EPSF staff generated versions of Draw A Person (EPSF.DAP), Preschool Language Scale (EPSF.PLS) and a Motor Activity Scale (EPSF.MAS). The EPSF.PLS and EPSF.MAS have five subtests and three subtests, respectively, labeled PLS I through PLS V and MAS I through MAS III.

PLS I through PSL V are described in the 1992 PLS-R manual as measures of "visual-vocal integration, vocabulary, auditory integrative responses, integrative auditory memory and discriminative visual-auditory memory", respectively, (Werner, 1992b, p. 2). The MAS I, MAS II and MAS III are discribed in the 1992 MAS manual as measures of "body imagery and spatial orientation in relationship to body parts; manual dexterity; and body control", respectively, (Werner ,1992a, p. 3).

The EPSF screening tests battery generates a developmental "diagnostic profile" (see Appendix A) composed of seven developmental areas including Receptive Language (RL), Expressive Language (EL), Auditory (AU), Visual Memory (VM) , Visual Discrimination (VD) , Fine Motor (FM) and Gross Motor (GM) . These seven developmental areas are generated from composites of selected EPSF screening tests and/or subtests.

The EPSF screening battery is critical in the identification process of individual children's developmental strengths and potential need areas. The screening battery generates a significant level of developmental strength or risk for each child through categorization in each of the seven generated modality areas. A child can be functioning in one of five different defined developmental risk level known as "considerable strength (CS), moderate strength (MS), average (AV), moderate need(MN) and considerable need (CN)". These aforementioned five developmental risk levels are equivalent to two years or more above developmental age expectancy (CS), one year above age expectancy (MS), at expected age level (AV), 1 year approximately below age expectancy (MN) and 2 years or more below age expectancy (CN), respectively (Werner, 1990; Roth, et al. 1993).

EPSF Summative Evaluation Studies of EPSF Treatment Program

Much of the previous research on the EPSF program has involved periodically generated EPSF summative evaluation reports as well as some independent research studies in the literature. The vast majority of previous EPSF summative research has focused on the effectiveness of the EPSF program with a few independent studies done on validation of the EPSF screening battery and resultant EPSF modalities. Werner (1990) simply states the EPSF "screening process is a valid and reliable process determining all children's developmental levels in receptive and expressive language, auditory, visual discrimination and memory, fine and gross motor modality areas" (p.15).

EPSF yearly and periodic summative evaluation summaries since the mid 1970's "appear impressive to those not versed in research methodology" (McConnell, 1986, p. 29). Strand and Werner (1981) in the EPSF evaluation summary of the EPSF project from 1971 to 1981, consistently focused on the use of the pre-post test design with gain scores on the PPVT, VMI and EPSF.PLS used as the dependent measures of the EPSF program effectiveness with no mention of the EPSF. MAS or EPSF.DAP as program effectiveness measures. Numerous subjects numbering over 1000 or more were mentioned but not documented in the annual EPSF project evaluations.

Strand and Werner (1981) presented mean gain scores on the EPSF.PLS, PPVT and or VMI for EPSF treatment program defined high risk participants were presented. Strand and Werner (1981) state for every month in the EPSF treatment program, children made 3.0, 2.0 and 1.65 developmental month gains on EPSF.PLS, PPVTand VMI post testing, respectively. Methodological limitations were noted including no use of comparison control groups with sole reliance on pre-post gain scores as program effectiveness measures.

McConnell (1986) talks in length about the serious flaws in using gain scores in a test-retest methodology despite the fact that "gains in achievement between pretest and posttest as a measure of effectiveness of a treatment makes good intuitive sense" (p. 30). Patrick, Kimball and Crawford (1984) in their meta-analysis of the 1971-1981 EPSF summative report stated the same critical comments regarding gain scores and the use of pre-post single group design. Subsequent discussion of the previous EPSF summative reports after 1981 continue to contain similar concern over research methodology (Bryant, 1991; Terbush, 1990; and Driscoll, 1992).

Major methodological concerns with gain scores include: (1) the phenomena of statistical regression to the mean, in effect, low scores increase and high scores tend to decrease upon retesting of the same subject; and (2) the difficulty of using the assumption that equal raw scores represent equal increments in achievement gain. "Initial high scorers on the pretest must pass the most difficult test items to increase their scores while initial low scorers can answer easier items on the post test and show relative large mean gain" generally on tests with ascending order of item difficulty like the EPSF PPVT-R and VMI tests (McConnell ,1986, p. 32). Also, (3) ceiling effects of high scorers who have little room for gain when retested. For example, McConnell (1986) states some of the EPSF students in her study made the maximal possible score on the initial PLS and MAS testing.

Other criticisms of the EPSF summative research focus on the lack of longitudinal studies looking at overall lasting effects of the EPSF treatment program. Strand and Werner in response to suggestions from annual EPSF reports did a initial longitudinal study dating 1979 thru 1982 involving EPSF treatment for defined moderate and high need kindergarten children compared to a control group both given the PPVT and VMI as initial study measures. Subjects were given the PPVT, VMI and Gates MacGinitie Reading Test as subsequent measures of EPSF program effectiveness.

Strand and Werner (1981) reported a one year gain in overall Gates MacGinitie Total Reading achievement for the experimental EPSF "moderate" and "high- risk" children versus the control group (Driscoll, 1992). However, the study was published in EPSF summative literature with limited discussion of the number of subjects in the study which was composed of "five pairs of experimental and control subjects for each school" but in "some cases less than five pairs were used" (Strand & Werner, 1981, p. 37). Other noted Strand and Werner (1981) study limitations include the use of only the Gates MacGinitie Total Reading score (and not Vocabulary and Comprehension subtests scores) data as well as the exclusion of the EPSF screening battery EPSF.PLS, EPSF.MAS and EPSF.DAP data (in effect, 60% of the battery) in the study.

Betz (1990) states a 1985 to 1988 EPSF educational program intervention summative report was done on the effectiveness of the EPSF-Success program designed for educationally at risk kindergarten and first grade students. The two groups initially in 1985 consisted of a total of 452 subjects (361 experimental and 91 control). The experimental group of defined "moderate" and "high-risk" kindergarten students received the EPSF kindergarten and first grade treatment. The control group received "regular kindergarten and first grade services" (Betz, 1990, p. 6). Stated factors reduced final sample size to 137 experimental and only 28 control students.

No statistical data was presented on the significance of the Betz (1990) study except that the experimental group "gained more (<.005) than the comparison group of 28 students" (p.7) on PPVT and IOWA Test of Basic Skills (ITBS) over the three year period. This is the same study used for the 1990 United States Department of Education JDRP/PEP national validation study of the EPSF-Success program for inclusion as an exemplary program in the

National Diffusion Network (NDN) according to Betz (1990).

Independent Studies of EPSF Treatment Program

Independent research on the effectiveness of the EPSF treatment program has been noted in the literature (Baenen,1992; Driscoll, 1992; McConnell, 1986; Parker & Clechalski, 1990; Patrick, et al. 1984; Roth,et al. 1993; Zeh & Baenen, 1991). Overall mixed results as to the long term effectiveness of the EPSF treatment program were noted. Difficulty comparing the effectiveness of the EPSF treatment program in the different research studies has been noted due to such a wide variety of independent variables used including such measures as special education placement, student grade retention, performance on achievement test measures such as the ITBS, California Achievement Test (CAT), Metropolitan Readiness Test, Wide Range Achievement Test, Comprehensive Test of Basic Skills (CTBS), and Stanford Achievement Test. Time intervals between dependent and independent measures varied from nine months to two years.

EPSF Screening Battery Research

Research by the EPSF project office on the EPSF screening battery has focused on the development of the EPSF staff generated MAS, PLS-R and EPSF.DAP instruments over the last twenty years. The VMI and PPVT have been well established as standardized screening instruments since their inclusion in the EPSF screening battery. Criticism over the years has focused on the need for extensive standardization of the MAS, PLS-R and EPSF.DAP (McConnell, 1986; Terbush, 1990; Bryant, 1991). In fact, some confusion was noted over substitution in the review of the literature in some EPSF studies (e. g., Terbush, 1990 and Driscoll, 1992) of the Preschool Language Scale (PLS) by Zimmerman developed in 1969 versus the EPSF.PLS developed by EPSF staff.

The EPSF Project Office conducted reliability studies on the EPSF.PLS, EPSF.MAS and EPSF.DAP during 1989 through 1992. Standardization and national norming of the MAS, PLS-R and EPSF.DAP occurred during 1987 and 1988 using the same sample of "3,093 children ages ranging from 4.6 to 6.6 from 42 different communities" (Werner, 1992a). Noted concern over the supposed national standardization sample is the fact that 34 of the 42 communities were in the North Central region of United States (with 21 of the 42 communities in state of Illinois).

The noted reliability studies of the MAS, PLS -R and EPSF.DAP involved 400 children drawn from the original standardization sample. Werner (1992a) states, regarding the reliability sample size, that "the relative small size of the sample is because of the large amount of data required of each student" (p.12). The noted reliability studies sample size for each of the six defined three month chronological age intervals for the PLS. MAS and EPSF.DAP range for 103 to 26 subjects.

Independent research on the EPSF screening battery has focused on four or five of the EPSF screening tests and/or their resultant EPSF modalities as predictor variables, (Agostin, 1993; Bryant, 1991; McConnell, 1986; Terbush, 1990; Roth, et al.1993). McConnell (1986) studied 116 kindergarten children using four of the five EPSF screening battery tests (excluding the EPSF.DAP) in her study. She was able to generate five of the seven EPSF defined modality areas. Experimental subjects received daily EPSF training with the control group receiving regular kindergarten instruction. Kindergarten school year end evaluation included the EPSF post testing, Metropolitan Reading Test, and teacher ratings as measures of kindergarten EPSF treatment program effectiveness.

McConnell (1986) found the PPVT-R, EPSF.PLS and VMI with equal weights explained 39% of the variance with EPSF.MAS adding little to the

EPSF screening battery predictor variable. No difference in adjusted means for the experimental and control groups were noted in tests of significance for EPSF treatment effects. Major limitations of the study included the limited sample size of 116, different school settings for the experimental and treatment groups and exclusion of the EPSF.DAP in the study.

Terbush (1990) studied the predictive validity of all five EPSF screening battery tests. He administered the EPSF screening battery in August, 1986 to 137 kindergarten children in two Arizona elementary schools with follow-up testing two years later in the Spring of 1988 with the Iowa Test of Basic Skills (ITBS). Significant predictive capabilities for the PPVT-R, EPSF.PLS, EPSF.DAP and VMI were found for determining overall (ITBS) test performance. The predictive correlations for the PPVT-R, VMI, EPSF.DAP and EPSF.PLS were found (using canonical variate loadings) to be .76, .69, .64 and .53 and determined to be statistically significant. The EPSF.MAS correlation of .37 was deemed insignificant as predicting future ITBS performance.

Terbush (1990) noted limitations of the study involved only students who scored high enough on the EPSF to be placed in regular kindergarten were included in the study. Thus, developmental kindergarten students were excluded in the Terbush (1990) study. Terbush, Bliss, Staines, Deneshinsky & Dankard (1990) in a follow-up study presentation at a national conference recommended longitudinal replication of the study.

Bryant (1991) studied 190 kindergarten students randomly selected from 26 elementary schools in Washoe County, Nevada. He compared the kindergarten generated seven EPSF screening modalities to a kindergarten year end locally developed kindergarten achievement test. Bryant (1991) found that the EPSF modalities of Receptive Language (RL), Auditory (AU) and Expressive Language (EL) provided the majority of the predictive validity for the local developed kindergarten achievement test with discriminant function correlations of .86, .77 and .69, respectively. (The EPSF modalities of VD, VM, FM and GM had discriminant function correlations of .36, .31, .29 and .25, respectively). The most obvious study limitation was comparing EPSF modality scores to a non-standardized kindergarten achievement test.

Roth, et al. (1993) used 161 kindergarten students for the combined 1985-1986 and 1986-1987 school years in a Maine school system. The purpose of the study was to determine if the 1985-1986 and 1986-1987 school year modality scores could predict whether kindergarten students would later be retained, referred to special education or placed in special education through the 1989-1990 school year. Roth, et al. (1993) found the EPSF Fine Motor (FM) modality was "consistently found (using stepwise discriminate analysis) to be a strong predictor of retention, referral or special education placement", (p. 357).

Roth, et al. (1993) used Tests of Basic Skills (ITBS) Total Reading scores as a secondary dependent measure with the best predictor (at the .05 significance level) of Fall ITBS reading achievement being the EPSF Auditory (AU) modality with a discriminant function factor loading of .22. The EPSF Fine Motor (FM) and Auditory (AU) modalities were found to be significant at the .05 level as the best predictors of Spring ITBS Total Reading scores with discriminant function factor loadings of .33 and .26, respectively. Overall, the EPSF screening battery was a significant statistical predictor of student status except in the modality area of Gross Motor (GM).

Agostin (1993) conducted a study of 184 children enrolled in kindergarten and first grade from three different elementary schools in the Fall of 1990. The subject pool included at risk children as defined by being second year kindergartners, first grade assisted or retained first students. Each student's kindergarten entry EPSF battery test results, Social Skills Rating System (SSRS) data and Stanford Achievement Test (SAT) results were gathered at the end of first grade. Research questions involved 1) which SSRT or EPSF modalities were the best predictors of SAT achievement at the end of first grade, and 2) which SSRT and/or EPSF modalities were the best discriminators among defined at risk children.

Agostin (1993) found the best predictor of SAT Total Reading considering both the SSRT and EPSF was found to be Receptive Language (RL) which predicted 39% of the variance. The best EPSF or SSRT predictors of SAT Total Math and Language were found to be Receptive Language (RL) and Visual Memory (VM). Overall, Agostin (1993) found that the RL modality accounted for from 23% to 24% of the total variance predicted by either the SSRT or EPSF in SAT Total Reading, Total Math or Language. Several limitations of the study were mentioned.

SUMMARY

The last twenty five years have seen a dramatic increase in the use of kindergarten screening in this country. This increase in the use of kindergarten screening has been due to various reported factors noted in the literature including (1) the growing body of research on the importance of the preschool years on later development of the individual. Also (2) the development of specific federal funded preschool programs such as Head Start and Follow Through and (3) the rapid growth of preschool and formal kindergarten education in the United States have been contributing factors to the need for kindergarten screening. The emphasis on kindergarten screening has been based in the literature on the assumptions that (1) early intervention for younger children can potentially decreasing the magnitude of any potential developmental problem and (2) reduce the cost factors of potential more extensive treatment or intervention, if the potential developmental problem goes undetected for a possible critical period of time.

The initial demand for kindergarten screening eventually lead to research in the literature on its limitations and best practices in use. The APA as early as 1974 published best practices for use with kindergarten and preschool screening procedures (McConnell, 1986). During the 1970's and 1980's research in the literature regarding overall kindergarten and preschool screening limitations included noting the unsophisticated test taking skills of preschoolers, typical heavy loading of verbal screening items with some young children having undeveloped verbal skills, test subject motivational errors, increased incidence of test rapport issues with preschoolers, premature labeling of children, limited generalization of results, difficulty handling the issue of rapid developmental changes in preschoolers as well as specific psychometric difficulties in preschool and kindergarten screening tests.

The rapid influx of kindergarten and preschool screening tests occurred during the 1970's. Some of the early major concerns in the literature regarding overall kindergarten screening were notably summarized by such researchers as Adelman and Lindsay & Wedell. Adelman as early as 1982 had made statements regarding widespread application of screening procedures such as "another example when pressure and enthusiasm for new procedures have lead to inappropriate extrapolation of research finds and premature applications' (p. 255). Lindsay and Wedell (1982) summed up the concerns of many researchers in the literature by stating "it is worrying when (screening) instruments are used up to 10 years with very little evidence of their usefulness" (p. 214). These previous concerns in the literature as to the technical merit of kindergarten screening tests are still relevant. Recently, Meisels (1989) stating that "developmental screening tests are in widespread use but few reliable and valid tests are available" (p. 578).

Specific concerns over the technical merit of kindergarten screening tests noted in the literature include such issues as subtest item gradients, subtest and total test internal consistency, test floors and ceilings, sample size, reporting norm sample as well as test utility and validity concerns. Complicating the issue of kindergarten screening development and use has been the confusion over the theoretical differences and applied uses of diagnostic, readiness and screening testing. Several researchers including Adelman (1982), Satz and Fletcher (1988), Meisels (1985,1987, 1989) and most recently, Gridley, Mucha and Hatfield (1995) have discussed the distinctions between screening, readiness and diagnostic tests.

Specific discussion in the literature on kindergarten screening reliability and validity issues was noted by previous researchers including Barnes (1982), Bracken (1987), Joiner (1977), Lehr , et al. (1987) and Meisels (1987) reporting the marginal statistical properties , especially predictive validity, of many screening tests. Meisels (1987) emphasized the seriousness of the need for predictive validity in screening tests by stating that screening tests without validity data is "an abuse of testing procedure and of the trust the community places in professional educators" (p. 6). Meisels (1985) had commented on screening tests having typically " marked decline in accuracy of prediction over a two year period or more " (p. 29). The importance of predictive validity of kindergarten screening tests and test batteries was noted in the literature by such notables as Bailey and Wolery (1989), Lindsay and Wedell (1982), Lichenstein and Ireton (1984) as well as Meisels (1987).

The review of the literature found kindergarten screening of academic skills, especially reading, has increased dramatically in the last decade as increased stress in formal academic learning for kindergarten children has occurred (Shephard & Smith, 1986; Charlesworth, 1989; and Slavin, et al. 1994). Slavin, et al. (1989 & 1994) and Charlesworth (1989) noted that increased emphasis since the 1980's on academic skill acquisition in earlier grades, especially kindergarten, coupled with increased use of kindergarten retention has lead to the increased development of developmental first and transitional first grade classes. Increased maladaptive use of kindergarten screening tests as academic readiness and diagnostic tests has been discussed in the literature (Meisels, 1987 and Satz & Fletcher, 1988).

Emphasis in the literature on academic skill acquisition, especially reading, has been documented since Durrell and Sullivan's research in the 1930's. Increased research emphasis has been noted of kindergarten screening for prediction of later academic achievement including significant studies by deHirsch, et al. (1966), Mercer, Algozzine and Trifiletti (1979), Horn and Packard (1985), Tramontana, Hooper & Selzer (1988) and Wallbrown, Engel, Wallbrown and Blaha (1975). Varied research results were noted in the literature regarding kindergarten screening test prediction of later academic achievement success due to various factors such as the academic measurement instruments used, the varied defined dependent and independent variables and the grade levels researched.

Various kindergarten screening programs were mentioned in the literature that have been developed to "get students off to a good start" (Slavin ,et al.1994). The Early Prevention of School Failure (EPSF) project is one of these kindergarten screening programs. The EPSF program has been designated by the United States Department of Education as an "exemplary program" (Educational Programs That Work, 1994). The EPSF program is based on six component parts (Werner, 1990). The initial EPSF program component is designated as "diagnosis" which reportedly looks at "the child"s developmental levels and preferred learning style" (Werner, 1990, p. 8). The basis of the initial EPSF "diagnostic" program component is the EPSF screening battery generated "diagnostic student profile" which generates 7 different developmental modality scores for the child. These scores and the EPSF overall 52 potential identified learning activities are used to generate the "curriculum design" component of the EPSF program (Werner, 1990).

The current study deals with the EPSF kindergarten screening battery of 5 tests, its 8 subtests ands the 7 derived developmental modality scores as predictors of end of first grade Gates MacGinitie Reading Achievement. The review of the literature found the majority of previous EPSF program research focused on the effectiveness of the EPSF school based intervention program. Previous independent research showed methodological flaws in the EPSF program developers' summary research reports (Bryan, 1991; Driscoll, 1992; McConnell, 1986; Patrick,et al. 1984; Terbush, 1990).

Previous EPSF screening battery predictive validity independent research noted in the literature was limited to only five studies (Agostin, 1993; Bryant, 1991; McConnell, 1986; Roth,et al. 1993; Terbush, 1990). Further predictive validity research on the EPSF screening battery for future reading achievement was needed,especially since the EPSF classroom based intervention program focuses on reading and math skill acquisition. No previous noted study in the literature could be found that studied the entire EPSF screening battery 5 tests, its 8 subtests and 7 developmental modality scores as valid predictors of future reading success.

Chapter III

Research Methodology

Subjects

The subjects for the present study were drawn from an annual subject pool of approximately 400 enrolled public school kindergarten students in a midwestern community of approximately 35,000 residents. The school system in the present study contains six elementary schools distributed throughout the community. The school system during the 1990-1991, 1991-1992 and 1992-1993 school years administered the EPSF screening battery as part of the school system's kindergarten enrollment policy.

During the 1990-1991 school year the designated school system in this study piloted the EPSF screening battery during summer kindergarten enrollment in four of the six elementary schools with subsequent use in the enrollment procedure of all six schools during the following 1991-1992 and 1992-1993 school years. The school system was able to allow kindergarten enrolled children to be administered the EPSF screening battery on several different occasions during the June through August prior to their formal kindergarten attendance.

The school system routinely administered the Gates MacGinitie Reading Test (Level 1) to first grade students in April or May of their first grade year. Approximately one half of the original sample of 1990 through 1992 EPSF screened children were not included in the study due to student attrition.

The current overall sample of subjects includes a total of 630 kindergarten enrollees who were identified from the 1990, 1991 and 1992 EPSF screening battery testing with corresponding follow-up Gates MacGinitie testing results available from the Spring of 1992, 1993 and 1994, respectively.

The current study includes two subject samples of 373 and 630 subjects. The 373 subjects sample is a subset of the overall 630 subjects. Thus, the current samples are similiar in age and sex ratio. The 630 subjects sample is comprised of 340 females (54% of the sample) and 290 males (46% of the sample) with an overall composite EPSF screening testing date average age of 66.62 months. The 373 subjects sample is comprised of 208 females (56% of the sample) and 165 females (44% of the sample) with an overall composite EPSF screening testing date age of 66. 68 months. The current sample includes children from all socioeconomic levels and numerous nationalities due, in part, to the location in the community of a well known university enrolling over 17,000 students. Race and sex were not used as study variables.

Instrumentation

The independent or predictor variables in this study include the EPSF screening battery including all five tests, the eight total MAS and PLS subtests and resultant seven developmental modalities. The dependent variables in this study are the Gates MacGinitie Reading Tests (Level 1, Form K) Total Reading, Comprehension and Vocabulary scores. Werner (1990) noted the EPSF screening battery is compromised of five tests including the Peabody Picture Vocabulary Test- Revised (PPVT-R), Developmental Test of Visual Motor Integration (VMI), EPSF.Preschool Language Scale (EPSF.PLS), EPSF.Motor Activity Scale (EPSF. MAS), EPSF.Draw A Person (EPSF. DAP). The EPSF. PLS has five subtests designated PLS I through PLS V and the EPSF.MAS has three subtests designated as MAS I, MAS II and MAS III.

The EPSF screening tests and subtests generate seven EPSF developmental modalities designated as Receptive Language (RL), Expressive Language (EL), Auditory (AU), Visual Discrimination (VD), Visual Memory (VM), Fine Motor (FM) and Gross Motor (GM). These EPSF generated modality areas as well as their corresponding standardized test or relevant subtest areas are as follows:

MODALITY AREA	INSTRUMENT
Receptive Language	PPVT-R, MAS I
Expressive Language	PLS I, II, IV
Auditory	PLS III, IV
Visual Memory	PLS V, VMI, DAP
Visual Discrimination	VMI, PLS V
Fine Motor	VMI, DAP, MAS II
Gross Motor	MAS III

PPVT-R

The PPVT-R (Dunn & Dunn, 1981) was originally developed in 1959 and revised in 1981 with noted use in hundreds of research studies. "It is a non - verbal multiple choice test designed to evaluate the hearing vocabulary or receptive knowledge of vocabulary in children and adults " (Sattler 1990, p. 348). Each child taking the test is "asked to point to one of four pictures on a page, then visually demonstrate the word the examiner has provided" (Terbush, 1990, p. 14). The PPVT-R takes 10 to 15 minutes to administer and does not require verbal responses of the test subject.

The PPVT-R was standardized on a representative national sample based on the 1970 U.S. Census of 4200 children ages 2. 5 through 18 and a selected sample of 828 adults ages 19 to 40. Split half reliability coefficients for the children ages 2.5 to 18 ranged from .67 to .88 with a .80 median. Alternative-forms reliabilities for a sample of 642 children ranged from .74 to .89 with a median of .81. A sample of 962 children given Forms L and M within a time interval of 9 to 31 days received alternative-from reliabilities ranging from .50 to .89 with a median of .76. The PPVT-R correlates .70 with the PPVT. McConnell (1986) states the PPVT-R was noted by Dunn and Dunn (1981) to have a "median correlation with other vocabulary tests of .71", (McConnell, 1986, p. 43) .

Sattler (1990) commented on the PPVT-R as "useful in measuring the extensiveness of receptive vocabulary "(p. 351) but cautioned it not to be used as a measure of intellectual functioning. He cautions special care should be used in the application of the PPVT-R with ethnic minority groups who tend to score lower on the PPVT-R than on intelligence tests potential due to their PPVT-R scores "may be in part a reflection of their verbal and experiential differences" (Sattler, 1990, p. 350). Tramontana , et al. (1988) stated, in their meta-analysis of 74 studies from 1973 to 1986, the PPVT was a " good predictor of reading at least in later grades" (p. 127).

Developmental Test of Visual Motor Integration

The Developmental Test of Visual Motor Integration (VMI), (Berry ,1989) is a "perceptual motor ability test for children ages 4 to 13" (Sattler, 1990, p. 368) used in the EPSF screening battery to measure visual discrimination and visual memory. The VMI was originally normed in 1964 and re-normed in 1981 with a standardization sample of 3,090 individuals ages 2-9 through 19-8. No information was provided by authors about the extent of VMI sample representation in relation to U. S. census data (Sattler, 1990).

The VMI instructions asks a child to copy a series of 24 presented increasingly difficult geometric shapes and forms from a simple straight vertical line ultimately to a complex six- sided star comprised of two double lined, overlayed triangles. The child's reproduced shapes are scored for accuracy by occurrence or omission of various types of errors such as rotations and detail expansion or constriction. Initial items are scored on pass-fail basis with overall quality of drawing scored on a 1 to 4 point weighted point system. Age equivalences, percentile ranks and standard scores are generated .

The VMI reliability and validity data in the manual is based on the original 1964 norms and data. VMI test-retest reliability coefficients for time intervals of

2 to 7 weeks ranges from .63 to .93 with a median of .81 (Sattler, 1990). Inter-rater reliabilities range from .58 to .99 with a median of .93 with internal consistency reliabilities ranging from .66 to .93 with a median of .79. Validity studies in the manual report concurrent validity being "satisfactory" (p."369) based on such criteria as chronological age (r. of .89), reading achievement (r. of .50), mental age (r. of .38 to .59), perceptual skills (r. of .80) and psycholinguistic skills (r. of .20 to .81), Sattler (1990).

Berry (1989) in his normative studies reported a "one year span internal consistency correlation ranging from .76 to .91 with median value of .85" (Driscoll, 1992, p. 44). Berry (1989) reported VMI correlations with other readiness tests averaging about .50 with correlation to reading higher for primary grades than for upper grades "with the tendency for the VMI to correlate more highly with arithmetic than with reading" (Driscoll, 1992, p. 44).

EPSF. Draw A Person

The EPSF.DAP is a human figure drawing task for 4.5 to 6.5 year old children used "as an estimate of a child's developmental cognitive ability to recall a meaningful image or impression and produce a graphic representation of a human form" (Werner, 1992a, p. 1). The EPSF.DAP is used in the EPSF screening battery to help create the Visual Memory (VM) and Fine Motor (FM) modality areas. Werner (1992a) in her EPSF.DAP manual states "while the Goodenough Harris Drawing Test is the model for the EPSF.DAP, the approach differs" (p. 1). She states the EPSF.DAP was specifically designed for children 4.5 to 6.5 to "measure a child's developmental of perceptual motor skills, concepts and strategies as well as recall an image (visual memory) normally expected to be associated with increases in chronological age" (Werner, 1992a, p. 1). Thus, the EPSF.DAP authors clarify that norms for other published DAP scoring systems are for older children.

The EPSF.DAP instructions require the examiner to say verbatim to the child: "On the page, I want you to make a picture of a person. Make the very best picture you can. Take your time and work very carefully" (Werner, 1992a. p. 20). The EPSF.DAP has a 28 point rating scale with 27 of the 28 points based on individual subjective criterion such as elaboration, proportionality. The last or 28th point on the EPSF.DAP rating scale is based on examiner's subjective view of the child's observed fine motor coordination. The EPSF.DAP score supposedly represents "an integrated measure of how the child perceives; uses his or her senses; grasps a tangible image from his or her mind and reproduces his or her concept of a person or image with pencil on paper" (Werner ,1992a, p. 1).

The EPSF.DAP 1992 manual states the standardization data was collected in September, 1990 on a national sample of 4,607 children ages 4.5 to 6.5 from 20 schools nationwide including 2,145 urban and 1,135 rural subjects. All students were reportedly from the regular classroom kindergarten population. The normative sample size for each of the six EPSF.DAP defined three month interval age groups ranged from 237 to 147. The EPSF.DAP manual states subjects in the normative sample were evaluated by teacher examiners who had been trained through the conventional two day EPSF workshop for EPSF program adopters. The "scoring accuracy" for the national sample was deemed in EPSF.DAP 1992 manual "not a concern due to the comprehensive (two day) training, number of raters (79) and size of sample" Werner, 1990, p. 2).

EPSF.DAP validity and reliability studies were reportedly done in 1990 and 1992 using a "representative sample" of 400 children ages 4.5 to 6.5 (80% being 5 year olds) from the original normative sample (Werner, 1992a, p. 14). Follow-up 1992 EPSF.DAP reliability research reported in the 1992 EPSF.DAP Manual found reported internal consistency estimates to be "slightly greater than 0.7", (Werner, 1992a, p. 14). The EPSF.DAP Manual noted EPSF 1992 internal consistency coefficients were "parallel" to as well as"were equivalent "in relationship to a 1987 Strommen & Smith Goodenough Harris DAP study (using a sample of 150 subjects, ages 5 through 8) "assessed by using the Kuder-Richardson Formula 20 reliability coefficient"' (Werner, 1992a, p. 14).

Reported 1990 EPSF.DAP construct validity with chronological age noted a correlation of .27 for boys and .22 for girls with the relatively low magnitude of correlations justified in the manual due to "a very limited age range in the EPSF sample (ages 4.5 to 6.5)" (Werner, 1992a, p. 16). Further construct validity evidence supposedly was illustrated in the 1992 EPSF.DAP manual through: 1) the pattern of most and least common features drawn by children in the validity sample with; 2) factor analysis presented on the 28 EPSF.DAP scoreable features. No statistical or specific data were noted in the EPSF.DAP 1992 manual to justify construct validity claims. Concurrent validity statements were made of "relative modest correlation between the DAP and the PLS, MAS and PPVT-R" with strongest relationship with the VMI" (Werner, 1992a, p. 17). Again, no specific statistical data is presented to justify validity claims.

Another noted limitation of the EPSF.DAP involves the manual reported 1992 reliability and validity sample of 400 subjects had only a range of 26 to 103 subjects in the six defined EPSF.DAP three month chronological age intervals from 54 to 77 months. Five of the six defined chronological age groups had less than 100 subjects with the youngest group (ages 54 to 57 months) and the oldest group (ages 74 to 77 months) have the fewest subjects, 26 and 29, respectively.

EPSF Motor Activity Scale

The EPSF Motor Activity Scale (EPSF.MAS) is described in the updated 1992 EPSF.MAS manual as "an instrument designed to assess a child's receptive language relevant to body parts and spatial orientation: manual dexterity and body control" (Werner, 1992b, p. iv). The test authors state that the EPSF.MAS is not a comprehensive assessment of perceptual motor development but is "designed to be used as a predictor of the developmental level of a child's fine motor and gross motor in relationship to his or her understanding of body language", (Werner, 1992b, p. 5).

The EPSF.MAS was originally developed in 1976 by Margot Heiniger and by EPSF staff. Bryant (1991) states that the EPSF.MAS was originally standardized on 138 subjects with geographic representation of the norming sample unknown and validity and reliability data not reported. McConnell (1986) stated that EPSF adopter school districts were not asked initially to report EFSF.DAP study scores to the EPSF Project Office with data on the EPSF.MAS "not included in any of the annual EPSF evaluation summaries" (p. 44).

The EPSF.MAS is comprised of three subtest categories noted as MAS I, MAS II and MAS III. MAS I is comprised of 13 items reported to measure" body imagery and spatial orientation in relationship to body parts" (Werner, 1992b, p. 3). The EPSF.MAS Manual states the MAS I has two sections noted A and B that reportedly measure (section A) "child's knowledge of body parts location and provide the examiner with a measure of language, body awareness and auditory memory", (Werner, 1992b, p.3) and (section B) "spatial orientation in relationship to body involving concepts such as over, under" (Werner, 1992b, p. 3). The child on MAS I items is asked to name eight body parts and hold a ball in five different spatial orientations to their body assessing the child's understanding of "under, front, top, between and right (versus left)", (Werner, 1992b, p. 3).

MAS II has only four items that proport to measure "manual dexterity" through stringing bead with both hands followed by with each hand: snapping fingers, finger to thumb touching and finger tip to opposite hand touching. MAS III contains 11 items designed to measure "body control" through activities of jumping, balancing on one foot, walking on preset pattern of footprints, skipping as well as tossing and catching yarn balls.

The examiner rates the child's EPSF.MAS performance of the 28 total subtest items with a possible maximum score of 30. The 28 items have remained the same since the test's inception. The tabulated test raw scores for each EPSF.MAS subtest area are entered into the EPSF computer input sheet where "MAS scores will be integrated with the outcomes of the total assessment process, not interpreted as a single measure of performance in any one developmental area" (Werner, 1992b, p. 16). The resultant MAS I, MAS II and MAS III areas are used to form components of the EPSF developmental modalities of Receptive Language (RL), Fine Motor (FM) and Gross Motor (GM), respectively. The child is then determined by EPSF computer program tabulations to be at one of five designated levels of modality developmental functioning ranging from "considerable strength" to "considerable need".

EPSF.MAS reliability studies noted in the EPSF.MAS Manual were done by Crawford (1989) and Thistlewaite and Cook (1992). Discussion of Thistlewaite and Cook's (1992) unpublished EPSF.MAS reliability and validity studies was done by Cook and Smith (1992). Crawford (1989) using an undescribed sample of children reported initial inter-rater reliability of .90 with a follow-up reliability coefficient after three months of .93. Thistlewaite and Cook (1992) with a "sample of 400 children from 10 school districts" (Werner, 1992b, p. 26) found internal consistency reliability coefficients for MAS I, MAS II and MAS III as .58, .60 and .60, respectively, using Cronbach's Alpha Formula. Sample size limitations comments noted with the EPSF.DAP are again relevant since researchers used the same subject pool for EPSF.MAS reliability studies.

"Validity of perceptual motor tests have been difficult to establish" (Werner, !992b, p. 24). Crawford (1989) reported comparing the EPSF.MAS to the Dayton Sensory Motor Survey and Purdue Perceptual-Motor Scale with the EPSF.MAS "concluded to be the best measure" (Werner, 1989, p. 34) No data was given in the EPSF literature or EPSF. MAS manual to support this claim. EPSF.MAS construct validity was reported in the 1992 EPSF.MAS Manual through: 1) the correlation between chronological age for boys of .27 and .21 for girls, and 2) the pattern of features found in "children's responses" (Werner, 1992b, p. 35) show a median item difficulty of 81% of items preformed correctly with 10 items being at or above the 90% correctly answered level. The 1992 EPSF.MAS Manual admits "relatively low MAS construct validity correlations with age but this can be ascribed to the very limited age range of EPSF sample (ages 4.5 to 6.4)" (Werner, 1992b, p. 35).

The EPSF.MAS was nationally standardized in 1988 reported "on a national random sampling based on the 1984 population data projected by the U.S. Census Bureau", (Werner,1992b, p. 35). A total sample of 3,093 children ranging in age from 4.5 to 6.5 were reportedly representative of ethnic, socioeconomic, community size, sex and age. The same standardization sample was used for the EPSF.PLS and EPSF.MAS. A total of 42 communities from five different regions of the U.S. were mentioned in the sample. The defined "North Central" region of the sample contains 1684 (over 54%) of total sample population with 33 (78 %) of all "national" sample communities located in Illinois or Ohio.

EPSF. Preschool Language Scale

The EPSF.PLS (PLS) was initially developed by the EPSF Project staff in 1971 due to the need for a test capable of measuring "integrated auditory-visual perception correlated to the typical preschoolers performance range" (Werner, 1992c, p. 1). The EPSF.PLS was then broadened to encompass cognitive processes in the areas of auditory, visual, kinesthetic and communicative language. Garner (1993) reported the stated purpose of the EPSF.PLS is "to predict school readiness by assessing integrated auditory, visual and motor synthesis" (p. 50) as noted from the PLS 1981 manual.

The EPSF.PLS is comprised of five subtests labeled PSL I through PLS V with a total of 50 possible points for 43 total different test items. The five subtests include:

1. PLS I Visual Vocal Integration

The child is asked to respond to examiner's oral individual questions to each of nine different stimulus pictures (e.g. "What is the mother doing to the bread" (slicing or cutting). McConnell (1986) stated PLS I is "regarded as more integrative than picture vocabulary tests in that both a statement and picture is presented to form a auditory-visual association requiring auditory and visual synthesis rather than rote memory", (p. 39). PLS I results are used with PLS II and PLS IV to form the EPSF Expressive Language (EL) developmental modality.

2. PLS II Vocabulary

The child is asked a series of eight questions without a visual clue that require the respondent to "demonstrate his/her understanding of the concept and not just make an association" (Werner, 1992c, p. 3). For example, "How does ice feel' (cold or wet). PLS II is used in conjunction with PLS I and PLS IV to construct Expressive language (EL) developmental modality.

3. PLS III Auditory

The child is presented with EPSF.PLS kit containing "familiar toys" (Werner, 1992c, p. 3) such as a doll and red block. The nine PLS III tasks range in difficulty level from one to three sequential directions for the child with the test focusing on "short term memory, association, sound discrimination and sequencing" (Werner, 1992c, p. 3). PLS level two difficulty tasks include "put two flowers and the doll in the box". PLS III is combined with PLS IV to generate the EPSF Auditory (AU) modality.

4. PLS IV Integrative Auditory Memory

This subtest is designed to assess grammatical closure and the child's ability to "recall stimuli received through his/her sense of hearing and based on his or her experience" (Werner, 1992c, P. 3). The child is presented with nine visually presented tasks with toy props accompanied by examiner verbally presented incomplete sentences to be completed by the child. For example, "this car is in a box. Now the car is (out)". PLS IV is used in both the EPSF Expressive Language (EL) and Auditory (AU) in combination with other PLS subtests.

5. PLS V Discriminative Visual-Auditory Memory

The child is presented with seven different tasks requiring him or her to visually remember and reproduce a series of briefly presented geometric shapes (triangle, circle or sun) involving four difficulty levels for recall from two to eventually four shapes. The PLS IV was the only PLS subtest where individual items are valued as 2 (versus 1) points per accurate item completion. Thus, the PLS IV subtest has a maximum of 14 possible points.

The initial EPSF.PLS normative sample consisted of 4,270 children from 37 school districts in nine states during the 1975-1976 school year, (McConnell, 1986). Werner (1992c) states "the PLS was standardized in a carefully selected national sample of more than 5,000 subjects in 1978 and more than 3,000 subjects in 1988" (p. iv). The 1988 EPSF.PLS national norming sample was the same sample used for the EPSF.MAS. Previous mentioned geographic representation limits of the sample include the fact that 34 of the 42 communities in the "national sample" were from the states of Illinois and Ohio.

The initial EPSF.PLS reliability studies in the 1981 PLS manual report that a test-retest correlation of .77 "in Summer 1973 with 97 pre-kindergarten children ages 4.5 to 5.5 in four Illinois communities" (McConnell, 1986, p. 41). Garner (1993) and McConnell (1986) reported the EPSF.PLS split-half reliabilities were not correlated in the 1973 reliability studies due to the subtests contained too few items for calculations to be done.

The 1988 and 1992 EPSF.PLS manuals mentioned Fredebaugh's (1984) study in the Virgin Islands done to establish the PLS test-retest reliability. Fredebaugh randomly selected 34 students from 17 EPSF classes with a 14 day test-retest interval. Overall reported test-retest reliabilities of .82 with PLS I through PLS IV reporting "strong correlations" but PLS I was the only PLS subtest "not highly correlated (r. 49)" (Werner, 1992, p. 29). The 1992 reliability study involved a sample of 400 children ages 4.5 to 6.5 (same sample as for EPSF.DAP and EPSF.MAS). The previous mentioned limitation of smaller sample size for five of the six derived sample age groups remains relevant (see

EPSF.PLS instrumentation narrative for further discussion). "Internal consistency of the PLS was computed using several formulas (Spearman-Brown, Cronbach's Alpha and Guttman's Split-Half). All of these reliability coefficients were on the order of 0.7" (Werner, 1992c, p. 29). Still no specific EPSF.PLS correlations are reported in either the 1988 or 1992 PLS manuals.

EPSF.PLS construct and criterion related validity studies were mentioned in the 1992 EPSF.PLS manual. The 1992 sample of 400 children generate a construct validity coefficient of .25 between the EPSF.PLS and chronological age. Patterns of children's EPSF.PLS responses were presented as a second proof of construct validity. The EPSF.PLS manual stated a median range of item difficulty of .50 with a range of .95 to .05. Factor analysis of children's EPSF.PLS responses suggested "some common pattern among children's responses" (Werner, 1992c, p. 32) with nine factors identified. Noted factor analysis found the first through fourth factors being parts of PLS V, PLS III, PLS II and PLS III & PLS IV, respectively. PLS I did "not group together in the factor structure" (Werner, 1992c, p. 32). Criterion validity studies mentioned in the EPSF.PLS manual reported EPSF.PLS total score correlations of .60, .42, .40 and .24 to the PPVT, VMI, EPSF.MAS and EPSF.DAP. No specific statistical data is given in the EPSF.PLS manual to support drawn conclusions.

Gates MacGinitie Reading Test
Third Edition

The Gates MacGinitie Reading Test has a long history of being a well known reading achievement measure and was a "prototype of the contemporary standardized reading test " (Calfee, 1985. p. 593). Lindquist (1982) stated "the Gates MacGinitie Reading Tests are designed not to be a diagnostic test but rather a survey of reading achievement" (p. 332). The Gates MacGinitie was first developed in 1926 with revisions in 1976 and 1989. The 1989 revised Gates MacGinitie contains nine levels "to assess student achievement in reading skills from kindergarten through grade 12" (MacGinitie

& MacGinitie, 1989c, p. 21). The test is divided into alternative forms K and L available for most grade levels. The test can be hand or machine scored.

MacGinitie and MacGinitie (1989b) reported their test raw scores are converted to standard scores reported in normal curve equivalent (NCE), percentile rank (PR), stanine, extended scale score (ESS) and grade equivalent (GE). Standard scores are generated on the Gates MacGinitie in the areas of Vocabulary, Comprehension and Total Reading Score. The Gates MacGinitie Total Reading score is basically a sum of the Gates MacGinitie Vocabulary and Comprehension scores. MacGinitie and MacGinitie (1989b) in their Gates MacGinitie Reading Tests (Third Edition) Manual for Scoring and Interpretation stated "the (Gates MacGinitie)Total (reading) raw scores are the sums of the Vocabulary and Comprehension raw scores", (p. 79).

The Gates MacGinitie Reading Test is comprised of a Vocabulary and Comprehension section with 45 and 48 test items, respectively. The Vocabulary and Comprehension areas require 20 to 25 and 30 to 35 minutes to complete. The Vocabulary subtest measures knowledge of words in isolation and evaluates" the student's knowledge of frequently used nouns, verbs, adjectives and other parts of speech" (Cooter & Curry, 1989, p. 256). Vocabulary items are in multiple choice format with four choices per test item. Each Vocabulary item has a visual clue. MacGinitie & MacGinitie (1989 a) stated regarding their vocabulary subtests that " levels 1 and 2 are primarily tests of decoding skills in which the child must sound out or recognize words that correspond to a picture", (p. 256).

The Comprehension subtest is a direct measure of the student's ability to read and comprehend. The Comprehension subtest is compromised of short one and two sentence reading passages followed by an implied question to answer or choose the best of three presented visual representation of the reading narrative. MacGinitie and MacGinitie (1989a) stated that on "Comprehension (subtest) levels 1 and 2, students begin with one sentence passages and must choose one picture that best reflects the meaning of the passage. As the test progresses , the narrative and expository passages become longer", (p. 257). Vocabulary and Comprehension responses are marked in the Gates MacGinitie test booklet.

The extensive national standardization of the Gates MacGinitie Reading Tests, Third Edition was done in the 1987-1988 school year involving "77,413 students in 222 schools in 67 school systems in 30 states" (MacGinitie & MacGinitie, 1989b, p. 25). Students in the sample were representative of the 1980 U. S. Census data regarding SES, school district size and region of the United States. Standardized data from the Fall 1987 and Spring 1988 testings of grades 1 through 12 was gathered with sample size per grade ranging for 1466 to 3589. An additional 25,210 students participated in the three "equating studies' to equate or statistically compare (1) the 1989 Gates MacGinitie alternative test forms, (2) two adjacent grade levels in Gates MacGinitie test responses and (3) Gates MacGinitie, Second Edition to the Third Edition. Overall the Gates MacGinitie, Third Edition had reliability coefficients for Vocabulary, Comprehension and Total Reading Score of .88 to .91, .87 to .92 and .93 to .95, respectively.

The current study involves the Gates MacGinitie, Third Edition (Level 1, Form K) with reliability coefficients (using the Kuder Richardson Formula 20) for the Vocabulary, Comprehension and Total Reading Scores of .93, .94 and .97. Extensive data on the Gates MacGinitie tests ceiling, floor and test completion rate was noted in the manual. Correlations were reported in the test manual of .88 between the Gates MacGinitie (Level 1, Form K) Vocabulary and Comprehension subtests with a .50 reliability coefficient of the differences between both Gates MacGinitie Reading Test (Level 1 , Form K) subtests (MacGinitie & MacGinitie, 1989c).

Reported validity studies in the 1989 Gates MacGinitie manual comparing Gates MacGinitie (Level 1) Vocabulary, Comprehension and Total Reading Scores to the Iowa Tests of Basic Skills , Comprehensive Tests Of Basic Skills, California Achievement Test, Metropolitan Achievement Tests, Survey of Basic Skills as well as to English course grades and grade point average found correlations ranging from .56 to .68, .83 to .88, .78 to .86, .45 to .72, .65 to .79, .77 to .83 and .68, respectively. Overall, the Gates MacGinitie Reading Test is reportedly adequate as a "gross first screening of reading ability" (Cooter & Curry, 1989, p. 258). MacGinitie and MacGinitie (1989c) stated the Gates MacGinitie Reading Test, Third Edition provides" an effective means of assessing general reading achievement", (p. 23). Overall, Cooter and Curry (1989), Graham (1990), Lindquist (1982) and MacGinitie and MacGinitie (1989a and 1989b) all reported the Gates MacGinitie to be suitable as a general screening of reading achievement. Thus, the Gates MacGinitie would definitely be beneficial as part of an academic screening program.

Design

This research is a correlational study using available archival data. There is no experimental treatment involved in this undertaking. This correlational study investigated the relationships between the EPSF Screening Battery tests, its subtests and the EPSF modality ratings as predictors of future performance on the Gates MacGinitie Reading Test, Third Edition (Level 1, Form K).

Procedures

The EPSF screening battery was routinely administered in a midwestern community of 35,000 to approximately 350 to 400 potential entering kindergarten children as part of the public school kindergarten enrollment policy during the 1990-1991, 1991-1992 and 1992-1993 school years. This screening was administered by school system screening teams consisting of teachers and other hired professionals. The screening team had received a two day EPSF training orientation. Screenings were done in the summer months prior to formal kindergarten attendance. Some kindergarten enrollees (44 children in this study) who missed the summer EPSF screening were given their EPSF screening battery during the first semester of Kindergarten. No formal widespread use of the EPSF treatment program was instigated in the school system despite the availability of EPSF classroom materials.

The school system in the current study had administered the Gates MacGinitie at the end of first grade as a screening instrument for potential Chapter 1 eligibility only in the four Chapter 1 target schools prior to the 1991-1992 school year. The Gates MacGinitie was adopted for use with all six elementary schools during the 1992-1993 school year. EPSF screening profile information for this study involved EPSF kindergarten screening profiles that could be matched to available first grade Gates MacGinitie data on the same students.

A total of 630 kindergarten student EPSF computer profiles with 93% gathered in the summer screenings of 1990, 1991, 1992 (having corresponding April of 1992, 1993 and 1994 first grade Gates MacGinitie scores) were collected with school permission from: 1) the EPSF computer data system records for all 630 subjects, and 2) individual EPSF raw data files for 373 subjects. The total age and raw score developmental age equivalences for the PPVT-R, VMI, and EPSF.DAP were collected from all 630 computer data sheets with total raw scores for the EPSF.PLS and EPSF.MAS for 371 subjects available. The EPSF computer sheet does not show the raw score data for the EPSF.PLS and EPSF.MAS subtests.

Each child's EPSF generated modality rankings for each of the seven developmental modalities were gathered from computer generated EPSF student developmental profile sheets. A numerical ranking of 1 was assigned for "considerable strength", a ranking of 2 assigned for "moderate strength, a ranking of 3 for "average", a ranking of 4 assigned to "moderate need" and a ranking if 5 assigned for "considerable need". Each of the student's seven individual developmental modality scores were thus assigned a numerical ranking of 1 to 5 for data analysis.

Data from the 630 Gates MacGinitie Reading Tests was obtained in Spring 1992, 1993, and 1994. The May 1992 Gates MacGinitie data was only available from the four school system Chapter 1 eligible schools. The 1992 and 1993 Gates MacGinitie data was available from all six elementary schools. The Gates MacGinitie Reading Test (Level 1, Form K) Total Reading, Vocabulary and Comprehension subtests scores are reported to the school district in the form of grade equivalencies, NCE, stanines and national percentile ranks.

Statistical analysis for predictive capabilities of the 5 EPSF tests, 8 subtests and 7 individual developmental modality scores were compared to the Gates MacGinitie Total Reading, Vocabulary and Comprehension results. An initial canonical statistical analysis was preformed on the composite Gates MacGinitie Vocabulary and Comprehension subtests. The Gates MacGinitie Total Reading Score was not used in the canonical analysis due to the fact it is a composite number of the two Gates MacGinitie subtests and potentially could weight the effects of the collapsed independent variable linear composite if entered into the canonical analysis. The Gates MacGinitie Total Reading Test score was used as an single dependent variable in the secondary multiple regression analysis preformed on the data.

Supplemental statistical comparison of the 135 EPSF defined "moderate need" and "considerable need" subjects' scores noted in this study to a random equal size sample of defined not at-risk EPSF subjects was not attempted in this study. Also, data analysis of sex and age of subject was not done in the current study.

Data Analysis

The defined predictive or independent variables in this study are the total scores of the 5 standardized EPSF Screening Battery tests, their 8 EPSF subtests and the EPSF defined 7 individual developmental modality scores. The 7 EPSF modality rating score were derived from the assigned value of a score from 1 (for "considerable strength") to 5 (for "considerable need") based on the 5 EPSF defined potential modality strength levels. Individual test subjects' age and sex were not used for secondary analysis as predictive variables.

The dependent or criterion set of variables involved the Vocabulary, Comprehension and Total Reading scores for the Gates MacGinitie Reading Test, Third Edition (Level 1, Form K) as measured by Normal Curve Equivalency (NCE). The use of NCE scores was done in this study versus potential use of the Gates MacGinitie Reading Test derived scores in the form of percentile ranks (PR) , grade equivalences (GE), stanines or extended scale scores (ESS). The use of NCE scores in this study was justified through relevant review of the literature and Gates MacGinitie test authors' comments.

The use of NCE derived scores for test result interpretation in this study was determined the most relevant Gates MacGinitie Reading Tests derived score. A brief review of the literature on derived scores is relevant. Frechtling (1989) reported that norm referenced tests offer a wide choice of derived scores but cautioned each has their strengths and weaknesses. Grade equivalences (GE) are noted in the literature to be the most easily misunderstood derived score (Frechtling, 1989; Gary, 1975; Green, 1987; Hanna, Dyck & Holen, 1980; Phillips & Clarizio, 1988 and Ward & Gould, 1980). Stanines are easily understood yet "provide a fairly gross measure of performance" (Frechtling, 1989, p. 477).

Percentile ranks (PR) are noted to "probably be the most widely used of the derived scores" (Green, 1987, p. 29) and are noted to be "easy to compute, universally used, applicable with a wide distribution of subjects and suitable for most test applications as one means of displaying information" (Brown, 1991, p. 346). Still, PR (1) suffer from being "time bound" in the sense they "are specific to the particular test, the particular reference group used, the time when the reference group was tested and the time when the school gives the test", (Green, 1987, p. 30), (2) should not be arithmetically averaged, (3) cause confusion that exist between PR and percentage correct, (4) should not be "submitted to any type of data analysis without converting it to some type of standard score" (Sattler, 1982 as noted by Brown, 1991, p. 26) due to not being an equal interval scale, distortion of scores can occur (especially as the ends of the score distribution) with, in effect, "equal percentile ranks definitely do not always represent equal differences in relative (distribution) position" (Brown, 1991, p. 25). PR are not recommended as valued data for comparative purposes in research (Brown, 1991; Frechtling, 1989; Green, 1987; MacGinitie & MacGinitie, 1989b and Rudner, 1989).

NCE are normal curve equivalencies and are "normalized transformations" of PR; therefore, giving the advantages of generating direct information about the relative status of an individual in a group with minimized distortion of scores and allowing for comparisons of individual and group scores at all points along the distribution. Three limitations of NCE were noted in the literature including (1) posing a potential communication problem to the lay person, (2) "can sound intimidating" (Flechtling, 1989. p. 477) and (2) some misinterpretation potential regarding NCE gain scores. Previous Chapter I Reading research (Talmadge, 1976) has noted that a zero NCE gain score means that the amount of learning was precisely what was expected. (In effect, a 50 NCE score always represents the exact average for that grade) Thus, some NCE gain score misinterpretations can occur by less knowledgeable test data interpreters.

MacGinitie and MacGinitie (1989b) in their Gates MacGinitie Reading Tests Manual for Scoring and Interpretation speak of how different derived scores can be used to answer different questions by data interpreters such as "How well does the child read ?", "As a group, how well do the children read ?" or " Has a new set of materials or procedures for teaching reading made any difference in how well the children can learn to read ? ". NCE was the only one of the derived score types used by MacGinitie and MacGinitie (including their own developed ESS derived score) mentioned by them to (1) answer all seven of their potential reading data interpretation questions and (2) have no noted limitations for their recommended Gates MacGinitie Reading test data reporting and interpretative use (MacGinitie & MacGinitie 1989b).

The study's initial level of statistical analysis required the use of a canonical correlation analysis (CCA) due to the fact that multiple dependent and independent variables were involved in this study (Pedhazur, 1982). CCA procedures allow the researcher to use larger number of criteria and predictor variables with a "reduction capacity similar to that of factor analysis" (Terbush, 1990). "Multiple regression, MANOVA, ANOVA and discriminant analysis can be shown to be special cases of canonical analysis" (Thompson, 1984, p. 7). CCA is a multiple regression technique that "is capable of showing the relationship between independent and dependent variables" (Thompson, 1984, p. 30). In effect, CCA analysis proceeds by initially collapsing each person's scores on the variables in each variable set into a single composite variable.

Three of the five previous EPSF predictive validity studies used the CCA technique at least as a portion of the basic study data analysis (Bryant, 1991; McConnell, 1986; Terbush, 1990). Some limitations in the interpretation of obtained CCA results are due to the fact that this statistical technique forms two linear composites (one involving the multiple dependent variables and one for the multiple independent variables) through the least squares analysis (Terbush, 1990). CCA is defined and limited by how the researcher generates the linear composites. The independent or predictor variables were divided in this study into three different sets of data for comparison to the dependent variables to answer each of the three research questions. The three different sets of predictor or independent variables included (1) the EPSF 5 screening battery tests, (2) the EPSF 8 screening battery subtests and the EPSF 5 screening tests and (3) the EPSF 7 developmental modality scores for each test subject.

The EPSF total scores for each of the five EPSF screening tests were combined to form one multiple independent variable linear composite for canonical comparison to the dependent variable linear composite composed from the Gates MacGinitie Reading Test (Vocabulary, and Comprehension scores) as measured by NCE derived scores. The second CCA data analysis involved comparison of a derived independent variable linear composite of the eight EPSF subtest raw scores and five EPSF screening test scores to the dependent variable linear composite of the Gates MacGinitie Reading Test (Vocabulary and Comprehension) scores as measured by NCE derived scores. The third CCA analysis involved a multiple independent linear composite of all seven EPSF modality scores compared to the dependent variable linear composite of the Gates MacGinitie Reading Test (Vocabulary, and Comprehension) scores as measured by NCE derived scores.

Specifically, the current study used CCA as the initial level of statistical analysis on each of the three research questions. These include the prediction of the Gates MacGinitie Reading Test (Level 1, Form K) Vocabulary and Comprehension scores through: 1) the degree of predictive variance accounted for by the five EPSF screening battery tests, 2) the degree of predictive variance accounted for by the five EPSF screening tests and eight EPSF subtests, and 3) the degree of predictive variance accounted for by the seven defined EPSF developmental modality areas. A Chi-Square test for statistical significance and statistical checks for nonlinearity were also conducted on the canonical analysis data.

Secondary statistical analysis of the current study individual research questions was done using the stepwise multiple regression statistical analysis technique. Agostin (1993) and McConnell (1986) used multiple regression statistical analysis in their noted EPSF screening battery predictive validity studies. Pedhazur (1982) stated that "basic multiple regression statistical analysis is eminently suited for analyzing the collective and separate effects of two or more independent variables on a dependent variable" (p. 6). Stepwise multiple regression is basically a variation of the forward selection multiple regression procedure. The forward selection multiple regression procedure is basically a multiple regression technique where ' the first predictor that has an opportunity to enter the prediction equation is the one with the largest simple correlation. If this predictor is significant, then the predictor with the largest semi-partial correlation with Y is considered, etc." (Steven, 1992, p. 87). Stepwise multiple regression is a varied form of the forward selection multiple regression method in that at each stage of the procedure a test is made of the least useful predictor. Thus the importance of each predictor is constantly being reassessed during the different stages of the stepwise multiple regression.

The majority of this study's statistical calculations were done for the canonical and stepwise multiple regression computations using the SSPS statistical package. The SPSS statistical package is noted by Pedhazur (1982) as " a versatile set of interrelated programs that afford great flexibility in data computation, data editing and data analysis" (p. 85). The SSPS statistical calculations for the current research were done through the Computing and Information Services of Oklahoma State University in Stillwater, Oklahoma.

The current use of the stepwise multiple regression was done to determine the best EPSF independent variable predictors of Gates MacGinitie Vocabulary, Comprehension and Total Reading scores as separate dependent variables. Each of the three study research questions required three separate stepwise multiple regression analyses involving the Gates MacGinitie Vocabulary, Comprehension and Total Reading scores as predicted by the three different independent variable sets of (1) the 5 EPSF basic screening tests, (2) the 5 EPSF screening tests plus the 8 EPSF subtest scores and (3) the 7 EPSF developmental modality scores.

Supplemental research question two stepwise multiple regression analysis was conducted using the defined independent variable subsets of derived scores from (1) the 8 EPSF subtests alone and (2) the 8 EPSF subtests and the PPVT-R, VMI and DAP screening tests thus, eliminating the PLS Total and MAS Total scores as predictors of future Gates MacGinitie Vocabulary, Comprehension and Total Reading scores.

Chapter 4

Results

Introduction

The purpose of this study was to examine the predictive validity capabilities of the basic EPSF Screening Battery 5 tests, its 8 subtests and derived 7 developmental modality scores as predictors of future reading achievement as measured by the Gates MacGinitie Reading Test (Form K, Level 1). Three research questions were addressed:

1. What is the degree to which the kindergarten age administered 5 EPSF screening battery tests are relate to and predict future, end of first grade, Gates MacGinitie Reading Test achievement?

2. What is the extent to which the kindergarten age administered EPSF screening battery 8 subtests are related to and contribute to the basic EPSF screening battery 5 tests' prediction of future, end of first grade, Gates MacGinitie reading achievement?

3. What is the degree to which the kindergarten age administered EPSF screening battery generated 7 individual developmental modality scores are related to and predict future, end of first grade, Gates MacGinitie reading achievement?

Statistical analysis of these three research questions was addressed through the initial use of canonical analysis followed by secondary analysis using stepwise multiple regression techniques. The majority of the statistical analysis calculations were performed using the SSPS statistical package locally available through the Computing and Information Services at Oklahoma

State University in Stillwater, Oklahoma.

This chapter is divided into four sections. The first three chapter sections each deal with a specific research question. Section four of the chapter is a summary of the major study results. Discussion of the results is presented in Chapter V.

Research Question One - What is the degree to which the kindergarten age administered EPSF screening battery 5 tests are related to and predict future, end of first grade, Gates MacGinitie reading achievement?

The first research question was initially studied using a canonical analysis of the overall relationship between the linear independent variables composite formed by the 5 EPSF screening battery test results in relationship to the linear dependent variable composite formed by the Gates MacGinitie Reading Vocabulary and Comprehension subtests results. The Gates MacGinitie Total Reading score was not used in forming the linear dependent variable due to it being a additive function or simple composite score of the Gates MacGinitie Vocabulary and Comprehension subtests scores. Follow-up stepwise multiple regression analysis was done to look at the capabilities of the independent variables to predict future Gates MacGinitie Vocabulary, Comprehension and Total Reading scores.

A canonical correlation was initially used to preform analysis of the relationship between the dependent and independent variables to eliminate the potential loss of valuable information from the variables caused by statistically looking at all them as separate entities. The canonical statistical analysis technique allows the researcher to state the relationships among variables more realistically - recognizing the fact that frequently in behavioral research variables are interrelated and not isolated functions. Also a check for non - linearity of the test data was preformed during the preliminary data analysis.

The present study involved the use of a sample of 373 subjects' data available for the computations necessary to answer research questions one and two. The larger sample of 630 subjects (including the 373 subjects from the sample subset) was used for research question three. Thus, at least two separate intercorrelation matrices had to be computed for the canonical and multiple regression statistical analysis of the three research questions. The 373 subject intercorrelational matrix is noted in Table 1. The 630 subject intercorrelational matrix in presented later in this chapter (see Table 19).

Table 1
Pearson Intercorrelational Matrix - 373 Subjects

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2	43														
3	47	94**													
4	19**	34**	31**												
5	13*	33**	35**	09											
6	16**	35**	33**	58**	17**										
7	11*	23**	22**	49**	02	66**									
8	03	18**	17**	52**	-00	62**	38**								
9	13**	31**	30**	40**	20**	64**	30**	32**							
10	04	20**	18**	46**	01	66**	49**	39**	33**						
11	15**	22**	19**	17**	22**	66**	20**	16**	20**	15**					
12	12**	15**	16**	31**	04	46**	37**	35**	30**	34**	22**				
13	12*	13*	13*	40**	-00	48**	43**	40**	34**	39**	14**	74**			
14	01	-04	- 02	04	03	09	08	06	04	06	06	16**	13*		
15	08	10*	11*	09	09	22**	18**	11*	15**	14**	15**	74**	20**	07	
16	02	09	08	10	07	05	03	04	07	00	01	10	07	02	09

1= Gates MacGinitie Vocabulary, 2 = Gates MacGinitie Comprehension, 3 = Gates MacGinitie Total Reading, 4 = PPVT-R, 5 = VMI, 6 = PLS Total, 7 = PLS 1, 8 = PLS 2, 9 = PLS 3, 10 = PLS 4, 11 = PLS 5, 12 = MAS Total, 13 = MAS 1, 14 = MAS 2, 15 = MAS 3, 16 = DAP

Note: Decimals omitted.
N = 373
** = p <01 * = p < .05

Table I shows the 5 EPSF tests of PPVT-R, VMI, EPSF.PLS Total, EPSF.MAS Total and DAP had three significant Pearson correlations among them. These three correlations significant at the .05 level were EPSF.PLS Total to PPVT-R, EPSF.PLS Total to VMI and MAS Total to EPSF.PLS Total with correlations of .58, .17 and .46, respectively (see Tables 1 and Appendix C). It is also interesting to note that the Table 1 matrix showed significant correlation between 4 of the 5 EPSF tests in relationship to the Gates MacGinitie Vocabulary, Comprehension and Total Reading Scores. The DAP was the only 1 of the 5 EPSF basic screening battery tests to not be significantly correlated with the Gates MacGinitie (see Appendix D for more specified view of data).

The canonical statistical analysis created synthetic Gates MacGinitie composite dependent variable (composed of the Vocabulary and Comprehension subtests) was statistically compared to the created artificial linear composite canonical independent variable (composed of the 5 EPSF screening battery tests). The SPSS statistical analysis yielded two canonical roots or variates. Only the first canonical root was statistically significant with a Chi Square of 101.203, df of 10 and $p < .001$. (see Table 2). Therefore, further statistical interpretation was only computed with Canonical Variate I. The canonical correlation squared (i.e. R squared) represents the "proportion of variance shared by a pair of canonical variates to which it corresponds", (Pedhuzur, 1982, p. 727). Thus, research question one Canonical Variate I produced 23% of the shared variance accounted for by the relationship between the linear composite of the 5 EPSF screening battery tests and the linear composite of the Gates MacGinitie Vocabulary and Comprehension subtests (see Table 2).

Table 2
Canonical Results of the 5 EPSF screening tests composite and the Gates MacGinitie Vocabulary and Comprehension subtests composite

Canonical Variate	Canonical Correlation	Canonical Correlation squared	Chi Square	df	p
I	.48360	.23387	101.203	10	<.001
II	.09253	.00856	3.164	4	>.500

Canonical variate structure coefficients or correlations (also known as loadings) were analysized to clarify the relationships defined by the first canonical variate. The canonical variate structure coefficients or canonical loadings represent the "correlations between the original variables and the composite variables" (McConnell, 1986, p. 66). Pedhuzur (1982) stated " as a rule of thumb, it is suggested that structure coefficients equal to and greater than .30 be treated as meaningful", (p. 732). Generally variables that are highly correlated with a canonical variate have more in common with it.

Standardized canonical coefficients or weights are used to determine the relative importance or contribution of variables. Pedhuzur (1982) stated that standardized canonical weights are analogous in interpretation to multiple regression beta weights and should be interpreted with caution due to their potential shortcomings. Canonical and multiple regression standardized function coefficients are typically used to generate a prediction equation to maximally predict individual scores on the specified variable being considered. (See Appendix F for canonical function coefficients or canonical weights). The current data analysis will focus on structure coefficient or factor loading discussion due to the increased score stability and less tendency for distortion factors, such as standard error.

Table 3

Canonical Variate Structure Coefficients of the 5 EPSF Test Scores and Gates MacGinitie Vocabulary and Comprehension Scores

Independent Variable		Dependent Variable	
PPVT-R	.71	Gates MacGinitie Vocabulary	.45
VMI	.69	Gates MacGinitie Comprehension	.99
PLS Total	.73		
MAS Total	.31		
DAP	.19		

N = 373

Table 3 illustrates that both dependent variables were significant structure coefficients or correlations but were more significantly loaded as a reading comprehension (versus vocabulary) measure. All the EPSF screening battery tests were significant canonical variate structure loadings except the DAP. The PPVT-R, VMI and PLS Total variate structure coefficients were highly significant and approximately equal loadings.

Redundancy coefficients were computed for the research question one derived canonical variate or root. "Redundancy coefficients are indexes of the average proportion of variance in the variables in one set that is reproducible from the variables in the other set", (Thompson, 1984, p. 25). The present research found that given the dependent variable linear canonical composite of Gates MacGinitie Vocabulary and Comprehension subtests scores, that 32.92% of the variance can be accounted for or explained by the independent variables linear canonical composite of the 5 EPSF screening battery tests. Also, given the independent variable linear canonical composite of the 5 EPSF screening battery tests, that approximately 60.28% of the variance can be accounted for or explained by the dependent variable canonical linear composite of the Gates MacGinitie Vocabulary and Comprehension subtests.

Secondary univariate statistical analysis of the dependent variables in relationship to the independent variables was done using the stepwise multiple regression technique. The inclusion of the Gates MacGinitie Total Reading score with the Gates MacGinitie Vocabulary and Comprehension subtests scores was done during this phase of the research question one data analysis. The generated Pearson intercorrelation matrix in Appendix F shows that all three Gates MacGinitie test scores are significantly correlated at the .05 level in both the 373 and 630 sample sets and that the Gates MacGinitie Total Reading and Comprehension subtest scores are highly correlated.

Table 4 presents the Betas (B), b weights, Y intercept, standard error (SE), t ratios and probabilities (p) for the independent variables in relationship to the Gates MacGinitie Vocabulary Subtest scores. The standardized regression coefficients (Betas) are used to assess the strength of the individual predictors. The strongest and only significant predictors of Gates MacGinitie Vocabulary scores were the PPVT-R and VMI. The standard error (SE) is used to set up a confidence interval around the predictor or independent variable scores. SE are typically desired to be smaller. The t-test of regression coefficients address the relationship between a given predictor and the criterion when other predictors have been taken into account.

Table 4

Betas, b-Weights, Standard Errors (S.E.), t Ratios and p Values for the Gates MacGinitie Vocabulary Scores as Predicted by the 5 EPSF Screening Battery Test

Variable	Beta(B)	b-Weight	S.E.	t Ratio	p Value
PPVT-R	.1540	.5501	.2233	2.464	.0142 *
VMI	.1134	.3493	.1591	2.195	.0288 *
PLSTotal	.0209	.1729	.5591	.309	.7573
MAS	.0592	1.0026	.9959	1.027	.3049
DAP	-.0120	-.0171	.0734	-.233	.8158

Y intercept - 37.5054

df = 367

* = p < .05

A stepwise multiple regression of the 5 EPSF screening battery tests as independent variables in predicting Gates MacGinitie Vocabulary scores was generated (see Table 5). The "R squared" presents the amount of accumulative variance accounted for by the specific independent variables entered by that step in the stepwise multiple regression. The "increase in R square" represents the increase in the amount of accumulative variance accounting for in the Gates MacGinitie Vocabulary scores as noted by the EPSF screening battery tests entered at that point in the stepwise multiple regression. The "adjusted R squared" or noted shrinkage represents the amount of variability that could be accounted for (at that point in the stepwise multiple regression) if the predictors were used with a different population. The "F equation value" and " F equation change" table discriptors represent with each step in the stepwise multiple regression, the exact F score and a quick visual representation of the change in F equation , respectively, at that step of the data analysis. The "significance of F change" represents a test of significance of the overall additive significance of the specific variables entered by that point in the stepwise multiple regression.

Table 5
Multiple Regression Results for Gates MacGinitie Vocabulary as Predicted by the 5 EPSF Screening Battery Tests

Step	Variable Entered	R Squared	Increase R Squared	Adjusted R squared	F Equation	F Equation Change	Sign. of F Change
1	PPVT-R	.0375	.0375	.0349	14.55	14.550	.000
2	VMI	.0508	.0133	.0457	9.90	5.182	.023
3	MAS Total	.0546	.0038	.0469	7.102	1.481	.224
4	PLS Total	.0549	.0003	.0446	5.340	.104	.747
5	DAP	.0550	.0001	.0421	4.272	.054	.816

N = 373

Table 5 shows (using R squared) that only 5.5 % of the total potential variance in Gates MacGinitie Vocabulary subtest scores can be accounted for by all 5 EPSF screening battery tests as the predictor variables and only approximately 4.2% (using the adjusted R squared) of the variance in a different population could be predicted. The only two statistically significant EPSF screening battery tests as predictors of future Gates MacGinitie Vocabulary were PPVT-R and VMI generated 92% of the total variance accounted for by the entire EPSF screening battery. The addition of the EPSF screening tests of PLS Total, MAS Total and DAP only added (to the PPVT-R and VMI) less than 1/2 of 1 % in Gates MacGinitie Vocabulary cumulative variance was accounted for by the EPSF screening battery.

The second stepwise multiple regression analysis used in research question one involved the Gates MacGinitie Comprehension subtest scores as predicted by the 5 EPSF screening battery tests. The generated Beta, b-weights, SE, t ratios and p values for this specific analysis are noted in Table 6. VMI, PPVT-R and PLS Total all were noted to be significant predictors of Gates MacGinitie Comprehension performance. Previous Pearson intercorrelation matrix data (see Appendix D) reported the VMI, PPVT-R and PLS similarly correlated to the Gates MacGinitie Comprehension subtest at the .33, .34 and .35 levels, respectively. Table 6 also shows the MAS Total and DAP were not significant EPSF screening battery predictors of Gates MacGinitie Comprehension reading achievement.

Table 6
Betas, b-Weights, Standard Errors (S.E.), t Ratios and p Values for Gates MacGinitie Comprehension Scores as Predicted by 5 EPSF Screening Battery Tests

Variable	Beta(B)	b-Weight	S.E.	t Ratio	p Value
VMI	.2797	.3528	.0587	6.009	.0000
PPVT-R	.2125	.3111	.0824	3.776	.0002
PLSTotal	.1890	.6405	.2063	3.105	.0020
DAP	.0428	.0251	.0271	.927	.3547
MAS	-.0201	-.1393	.3600	-.387	.6989

Y intercept - 11.0560
df = 367

The stepwise multiple regression summary table for Gates MacGinitie Comprehension scores (see Table 7) shows that 23.38 % of the variance in Gates MacGinitie Comprehension can be accounted for by the accumulative 5 EPSF screening battery tests. The results also show that the PLS Total, VMI and PPVT-R alone account for 23.18% of the variance in Gates MacGinitie Comprehension scores. The DAP and MAS Total contribute together only .2 of 1 % of the total variance accounted for by the overall 5 EPSF screening battery tests.

Table 7
Multiple Regression Table for Gates MacGinitie Comprehension as Predicted by the 5 EPSF Screening Battery Tests

Step	Variable Entered	R Squared	Increased R Squared	Adjusted R Squared	F Equation	F Equation Change	Sign of F Change
1	PLS Total	.1236	.1236	.1213	52.340	52.340	.000
2	VMI	.2007	.0771	.1964	46.453	35.674	.000
3	PPVT-R	.2318	.0371	.2255	37.109	14.924	.000
4	DAP	.2334	.0071	.2251	28.018	.804	.370
5	MAS Total	.2338	.0003	.2233	22.393	.105	.699

N = 373

Thus, 99% of the reported EPSF screening battery tests variance that contributes to the prediction of Gates MacGinitie Comprehension can be attributed to the PLS Total, VMI and PPVT-R. Also, the shrinkage in amount of variance when the same 5 EPSF screening battery tests are used to predict the Gates MacGinitie Comprehension scores applied to a different sample population was only to 22.33% (as compared to the 23.38% total variance for the original sample).

Table 8
Betas, b-Weights, Standard Errors (S.E.), t Ratios and p Values for Gates MacGinitie Total Reading Scores as Predicted by the 5 EPSF Screening Battery Tests

Variable	Beta(B)	b-Weight	S.E.	t Ratio	p Value
VMI	.3078	.3871	.0589	6.574	.0000
PPVT-R	.1893	.2763	.0827	3.345	.0009
PLS Total	.1570	.5302	.2068	2.563	.0108
DAP	.0304	.0178	.0272	.654	.5137
MAS Total	.0142	.0980	.3610	.271	.7862
Y intercept - 11.0560					

df = 367

Finally, the third portion of the stepwise multiple regression analysis for research question one deals with the prediction of the Gates MacGinitie Total Reading scores by the 5 EPSF screening battery tests. Table 8 shows that VMI, PPVT-R and PLS Total were all significant at the .01 level as individual predictors of the Gates MacGinitie Total score. The MAS and DAP were found not to be significant individual predictors of the Gates MacGinitie Total score.

Table 9 shows that overall 22.46 % of the variance in Gates MacGinitie Total Reading scores can be predicted by the accumulative 5 EPSF screening battery test with 21.40 % of the total variance still predicted by the overall 5 EPSF screening battery tests when the same prediction equation was used with a different population. Thus, less than a 5% change (from 22.46% to 21.40%) in EPSF screening battery predictive capabilities of Gates MacGinitie Total Reading scores was found when variance shrinkage was considered.

Table 9
Multiple Regression Results for Gates MacGinitie Total Reading Scores as Predicted by the 5 EPSF Screening Battery Tests

Step	Variable Entered	R Squared	Increased R Squared	Adjusted R Squared	F Equation	F Equation Change	Sign. of F Change
1	VMI	.1254	.1254	.1231	53.213	53.213	.000
2	PPVT-R	.2063	.0809	.2020	48.088	37.698	.000
3	PLS Total	.2235	.0172	.2172	35.399	8.160	.005
4	DAP	.2245	.0010	.2160	26.627	.464	.496
5	MAS Total	.2246	.0002	.2140	21.263	.074	.786

N = 373

The cumulative VMI, PPVT-R and PLS Total were found to be significant at the .01 level of significance as predictors of future Gates MacGinitie Total Reading achievement. The VMI, PPVT-R and PLS Total tests were found to all be significant at the .01 level as Gates MacGinitie Total Reading score combined predictors for Gates MacGinitie Total Reading scores. The VMI and PPVT-R alone were noted to be the most significant predictors and accounted for 20.63 % of variance of the Gates MacGinitie Reading Total scores. The PLS Total was noted to significantly add 1.72 % to the overall significant variance accounted by the VMI, PPVT-R and PLS Total. It is interesting to note that the initial intercorrelation matrix (see Appendix D) showed the PPVT-R, VMI and PLS Total all to have similar correlations of .31, .35 and .33, respectively, with the Gates Total Reading. The MAS Total and DAP were noted to add only .12 of 1% of the overall variance in the Gates MacGinitie Total Reading score accounted for by the EPSF screening battery.

Research Question Two -What is the extent to which the kindergarten age administered EPSF screening battery 8 subtests are related to and contribute to the EPSF basic screening battery 5 tests' prediction of future, end of first grade, Gates MacGinitie Reading achievement?

The second research question was initially studied through the use of canonical procedures to investigate the relationships between a linear composite formed by the independent variable and a linear composite formed by the dependent variable (composed of the Gates MacGinitie vocabulary and Comprehension subtests results). The independent variable canonical linear composite for the second research question added the 8 EPSF subtests to the 5 original EPSF screening test to determine their combined predictive capabilities. The 8 EPSF subtests include the five subtests of the PLS Total and the three subtests of the MAS Total.

An abbreviated intercorrelational matrix was generated for quick visual display of the EPSF subtests in relationship to the Gates MacGinitie Vocabulary, Comprehension and Total Reading scores (see Appendix G). Also, a check for non-linearity of the study data for research question two was done during the preliminary data analysis phase. Comparison of Appendix D and Appendix G shows that the 5 PLS subtests and the overall PLS Total test score on the basic EPSF screening battery all were significantly correlated with Gates MacGinitie Comprehension and Total Reading scores. Also, Appendix D and Appendix G show PLS I, PLS III ,PLS V, MAS Total (the EPSF basic MAS screening test) and MAS I were all significantly correlated with all three Gates MacGinitie reading scores.

The overall MAS Total basic EPSF screening test as well as MAS I and MAS III were all significantly correlated with Gates MacGinitie Comprehension and Total Reading scores. MAS II exhibited no significant correlation with any of the three possible Gates MacGinitie reading achievement scores. Overall, the PLS Total, PLS subtests, MAS Total and MAS subtests scores were less correlated with the Gates MacGinitie Vocabulary subtest scores than with either the Gates MacGinitie Comprehension or Total Reading scores.

A canonical analysis was done of the synthetic linear composite independent variable composed of the combined 5 EPSF tests and 8 EPSF subtests. The SPSS statistical package generated two canonical variates, or roots, of which only the first variate (designated Variate I) was deemed significant with a Chi Square of 120.79, df of 26 and p of < .001 (see Table 10). It was noted in Canonical Variate I that the dependent and independent variable composites had a correlation of .51. Canonical Variate I was noted to have 25. 85% of the shared variance between the dependent and independent variable linear composites. In comparison, the research question canonical variate I (involving only the 5 EPSF tests in the dependent linear canonical) had 23.38 % of the shared variance between the dependent and independent variables linear composites.

Table 10
Canonical Analysis Results of the 5 EPSF screening tests and 8 EPSF subtests composite and the Gates MacGinitie Vocabulary and Comprehension subtests composite

Canonical Variate	Canonical Correlation	Canonical Correlation Squared	Chi Square	df	p
I	.50841	.25848	120.079	26	<.001
II	.17428	.03037	11.226	12	>.500

Table 11 shows the calculated canonical variate I obtained variate structure coefficients or correlations (i. e. loadings). The standardized canonical function coefficients (or weights) are presented in Appendix H. The most significant independent variate canonical loadings were PPVT-R, PLS Total, VMI and PLS III with significance levels in the .60's. PLS I, PLS II, PLS IV, PLS V and MAS Total were significant but with loadings in the .30 to .45 range. The dependent canonical linear composite found Gates MacGinitie Comprehension to be the most significant factor loading, by far, with a value of .99. The noted canonical weights also found the Gates MacGinitie Comprehension to be the most significant dependent variable.

Table 11

Canonical Variate Structure Coefficients of the 5 EPSF Screening Battery Tests and 8 EPSF Subtests Scores and the Gates MacGinitie Vocabulary and Comprehension Subtest Scores

Independent Variables		Dependent Variables	
PPVT-R	.68	Gates MacGinitie Vocabulary	.47
VMI	.65	Gates MacGinitie Comprehension	.99
PLS Total	.69		
PLS I	.45		
PLS II	.34		
PLS III	.61		
PLS IV	.40		
PLS V	.44		
MAS Total	.30		
MAS I	.25		
MAS II	-.07		
MAS III	.21		
DAP	.18		

Redundancy coefficients were computed for the obtained significant canonical variate I. The current study found given the dependent variable linear composite of Gates MacGinitie Reading and Comprehension, that 20.32 % of the variance can be accounted for or explained by in the linear composite independent variable (comprised of the 5 EPSF screening tests and 8 EPSF subtests). Also, given the independent variable linear composite, that approximately 60.82% of the variance can be accounted for or explained by the dependent variable linear composite.

Secondary data analysis for research question two was accomplished using stepwise multiple regression techniques that allowed inclusion of the

Gates MacGinitie Total Reading score as a third dependent variable. Thus, adding data analysis potential to the available Gates MacGinitie Vocabulary and Comprehension subtest scores. The stepwise multiple regression analysis in research question two dealt with statistical analysis of the independent variable predictive capabilities for the Gates MacGinitie Vocabulary, Comprehension and Total Reading scores. The initial stepwise multiple regression analysis was of the combined 5 EPSF screening battery tests and 8 EPSF subtests (as the defined independent variables) predicting the Gates MacGinitie Vocabulary subtest results.

Table 12 shows that the only significant predictor of future Gates MacGinitie Vocabulary subtest achievement was the PPVT-R. The PPVT-R was noted to be at the .01 level of significance yet have a lower correlation of .22. Still the PPVT-R was noted to overall have a much smaller standard error (SE) than the vast majority of other potential EPSF predictors of Gates MacGinitie Vocabulary subtest achievement. Thus, Table 12 shows the PPVT-R was ,by far, the most effective predictor of first grade Gates MacGinitie Vocabulary subtest results (when each of the defined potential 13 EPSF independent variables including all 8 EPSF subtests and all 5 EPSF screening tests were statistically considered independently).

Table 12
Betas, b-weights, Standard Errors (S.E.), t ratios and p values for the Gates MacGinitie Vocabulary Scores as predicted by the 5 EPSF screening Battery Tests and 8 EPSF subtests

Variable	Beta(B)	b-Weight	S.E.	t Ratio	p Value
PPVT-R	.2212	.7906	.2424	3.262	.0012
VMI	.0789	.2430	.1646	1.477	.1407
PLS Total I	- 1.5387	12.7255	18.8315	-.676	.4996
PLS I	.4059	13.3413	18.8987	.706	.4807
PLS II	.2497	8.5163	18.9723	.449	.6538
PLS III	.5144	13.9398	18.6985	.746	.4565

Table 12 (continued)
Betas, b-weights, Standard Error (S.E.), t ratios and p values for the Gates MacGinitie Vocabulary Scores as predicted by the 5 EPSF screening Battery Tests and 8 EPSF subtests

Variable	Beta(B)	b-Weight	S.E.	t Ratio	p Value
PLS IV	.3485	10.4120	18.9531	.549	.5831
PLS V	.8151	14.5628	18.9767	.767	.4433
MAS Total	.0751	1.2705	2.9635	.429	.6684
MAS I	1.0036	.1089	3.6582	.030	.9763
MAS II	- .0059	- .1623	1.4190	-.114	.9090
MAS III	- .0048	- .1234	3.0744	- .040	.9680
DAP	- .0152	- .0217	.0734	- .296	.7672
Y intercept - 40.6325					

df = 359

Table 13 shows that overall that only 8.0 % of the variance in the Gates MacGinitie Vocabulary subtest can be accounted for by the 13 EPSF independent variables (consisting of the 5 EPSF screening tests and 8 EPSF subtests). Table 13 shows that the PPVT-R, PLS V and PLS II are the only significant independent predictors and account for 6.19 % of the variance in the gates MacGinitie Vocabulary score. Thus, 77. 37 % of the potential variance that could be accounted for by the independent variables was done so by 3 of the potential 13 independent variables. Thus, the other 10 EPSF independent variables only added 1.81 % of the accounted variance in the Gates MacGinitie Vocabulary subtest scores.

Table 13
Multiple Regression Results for the Gates MacGinitie Vocabulary as Predicted by the 5 EPSF Screening Battery Tests and 8 EPSF Subtests

Step	Variable Entered	R Squared	Increase R Squared	Adjusted R squared	F Equation	F Equation Change	Sign. of F Change
1	PPVT-R	.0375	.0375	.0349	14.555	14.550	.000
2	PLS 5	.0521	.0146	.0469	10.160	5.682	.018
3	PLS 2	.0619	.0098	.0543	8.115	3.868	.050
4	VMI	.0690	.0071	.0589	6.816	2.800	.095

Table 13 (continued)
Multiple Regression Results for the Gates MacGinitie Vocabulary
as Predicted by the 5 EPSF Screening Battery Tests and 8 EPSF Subtests

Step	Variable Entered	R Squared	Increase R Squared	Adjusted R squared	F Equation	F Equation Change	Sign. of F Change
5	MAS Total	.0733	.0044	.0607	5.808	1.723	.190
6	PLS 4	.0763	.0030	.0612	5.042	1.196	.275
7	PLS 3	.0783	.0019	.0606	4.429	.767	.382
8	PLS 1	.0786	.0003	.0583	3.879	.106	.745
9	PLSTotal	.0798	.0012	.0569	3.496	.477	.490
10	DAP	.0800	.0002	.0546	3.147	.090	.764
11	MAS 2	.0800	.0000	.0520	2.855	.012	.912
12	MAS 3	.0800	.0000	.0494	2.610	.009	.925
13	MAS 1	.0800	.0000	.0467	2.403	.001	.976

N = 373

The second stepwise multiple regression analysis for research question two found 8 of the 13 independent variables as significant predictors of future Gates MacGinitie Comprehension subtest achievement (see Table 14). These 8 independent variables in descending order of significance were VMI, PPVT-R, PLS III, PLS I, PLS V, PLS IV ,PLS Total and PLS II.

Table 14
Betas, b-Weights, Standard Errors (S.E.), t Ratios and p Values
for Gates MacGinitie Comprehension Scores as predicted by
the 5 EPSF screening Battery Tests and 8 EPSF subtests

Variable	Beta(B)	b-Weight	S.E.	t Ratio	p Value
PLSTotal	- 4.4487	- 15 .0731	6.9280	- 2.176	.0302
VMI	.2550	.3218	.0605	5.313	.0000
PPVT-R	.2439	.3570	.0892	4.005	.0001
PLS 3	1.4836	16.4702	6.8791	2.394	.0172
MAS 1	-.0951	-1.1667	1.3458	- .867	.3865
MAS 2	-.0536	- .6034	.5220	- 1.156	.2486
PLS 2	1.0586	14.7926	6.9799	2.119	.0347
DAP	.0409	.0240	.0270	.888	.3753
MAS Total	.0486	.3367	1.0902	.309	.7576
PLS 1	1.1844	15.9506	6.9528	2.294	.0224
PLS 5	2.1592	15.8055	6.9814	2.264	.0242

Table 14 (continued)
Betas, b-Weights, Standard Errors (S.E.), t Ratios and p Values for Gates MacGinitie Comprehension Scores as predicted by the 5 EPSF screening Battery Tests and 8 EPSF subtests

Variable	Beta(B)	b-Weight	S.E.	t Ratio	p Value
PLS 4	1.2627	15.4556	6.9728	2.217	.0273
MAS 3	-.0067	-.0667	1.1311	-.059	.9530
Y intercept	- 8.454655				

df = 359

Table 15 shows that overall 25.82 % of the variance of the Gates MacGinitie Comprehension subtest score can be accounted for by the 5 EPSF screening tests and 8 EPSF subtests. The results show that PLS Total, VMI and PPVT-R account for 23.18 % of the variance of the Gates MacGinitie Comprehension subtest that was predicted by total 13 defined EPSF independent variables.

Table 15
Multiple Regression Results for the Gates MacGinitie Comprehension Subtest as Predicted by the 5 EPSF Screening Battery Tests and 8 EPSF Subtests

Step	Variable Entered	R Squared	Increase R Squared	Adjusted R squared	F Equation	F Equation Change	Sign. of F Change
1	PLSTotal	.1236	.1236	.1213	52.340	52.340	.000
2	VMI	.2007	.0771	.1964	46.453	35.674	.000
3	PPVT-R	.2318	.0311	.2255	37.109	14.924	.000
4	PLS 3	.2368	.0050	.2285	28.547	2.429	.120
5	MAS 1	.2398	.0030	.2294	23.154	1.445	.230
6	MAS 2	.2422	.0024	.2298	19.494	1.149	.284
7	PLS 2	.2445	.0023	.2300	16.872	1.105	.294
8	DAP	.2461	.0017	.2296	14.856	.807	.369
9	MASTotal	.2471	.0009	.2284	13.236	.451	.502
10	PLS 1	.2474	.0003	.2266	11.898	.142	.706
11	PLS 5	.2480	.0006	.2251	10.823	.303	.582
12	PLS 4	.2582	.0102	.2334	10.441	4.934	.057
13	MAS 3	.2582	.0000	.2313	9.611	.003	.953

N = 373

The third portion of the stepwise multiple regression analysis involved the prediction of the Gates MacGinitie Total Reading score by the independent variables comprised of the 5 EPSF screening tests and 8 EPSF subtests (including PLS I through PLS V and MAS I through MAS III). Table 16 shows that VMI and PPVT-R are significant predictors (at the .01 level) of Gates MacGinitie Total Reading scores. The PLS I, PLS III and PLS V subtests were also noted to be significant predictors (at the .05 level) of future Gates MacGinitie Total Reading achievement.

Table 16
Betas, b-Weights, Standard Errors (S.E.), t Ratios and p Values for the Gates MacGinitie Total Reading Scores as Predicted by the 5 EPSF Screening Battery Tests and 8 EPSF Subtests

Variable	Beta(B)	b-Weight	S.E.	t Ratio	p Value
VMI	.3667	.2918	.0609	6.023	.0000
PPVT-R	.2954	.2024	.0897	3.293	.0011
PLS 3	14.8333	1.3404	6.9214	2.143	.0328
PLS 1	14.5432	1.0833	6.9955	2.079	.0383
PLS 5	13.9041	1.9054	7.0243	1.979	.0485
MAS 3	-.3311	-.0314	1.1380	-.291	.7712
MAS 1	-1.3298	-.1088	1.3541	-.982	.3267
MAS Total	.7255	.1050	1.0969	.661	.5088
MAS 2	-.3709	-.0331	.5253	-.706	.4805
DAP	.0164	.0281	.0272	.605	.5453
PLS 4	13.6523	1.1189	7.0156	1.946	.0524
PLS Total	-13.4854	-3.9926	6.9706	-1.935	.0538
PLS 2	13.4349	.9645	7.0227	1.913	.0565

Y intercept - 12.4632

df = 359

Table 17 shows that overall 24.4 % of the variance on the Gates

MacGinitie Total Reading score can be predicted by the 5 EPSF screening tests and 8 EPSF subtests (as compared to 22.46 % of the variance on the Gates MacGinitie Total Reading score can be predicted by the 5 EPSF tests alone). Also, Table17 shows that 22.43 % of the variance in the Gates MacGinitie Total Reading score can be accounted for by the VMI, PPVT-R and PLS III. These 3 independent EPSF variables were all significant at the .01 level with none of the other 10 EPSF independent variables noted to be significant. The other 10 EPSF independent variables only contributed 2.0 % of the 24.43 % total variance accounted for by all 13 EPSF independent variables. Thus, approximately 92% of the total variance predicted by the EPSF independent variables was contributed to by the VMI, PPVT-R and PLS III independent variables.

Table 17
Multiple Regression Results for Gates MacGinitie Total Reading
as Predicted by the 5 EPSF Screening Battery Tests and 8 EPSF Subtests

Step	Variable Entered	R Squared	Increase R Squared	Adjusted R squared	F Equation	F Equation Change	Sign. of F Change
1	VMI	.1254	.1254	.1231	53.213	53.213	.000
2	PPVT-R	.2063	.0809	.2020	48.088	37.698	.000
3	PLS 3	.2243	.0180	.2180	35.570	8.567	.004
4	PLS 1	.2293	.0050	.2209	27.368	2.368	.125
5	PLS 5	.2317	.0024	.2212	22.131	1.140	.286
6	MAS 3	.2326	.0009	.2200	18.487	.439	.508
7	MAS 1	.2337	.0011	.2190	15.904	.542	.462
8	MAS Total	.2346	.0009	.2178	13.948	.433	.511
9	MAS 2	.2357	.0011	.2167	12.438	.506	.477
10	DAP	.2363	.0006	.2152	11.203	.305	.581
11	PLS 4	.2364	.0001	.2132	10.162	.041	.839
12	PLS Total	.2366	.0002	.2112	9.299	.089	.766
13	PLS 2	.2443	.0077	.2170	8.928	3.660	.057

N = 373

Supplementary research question two multiple regression data analysis was preformed using (1) 11 EPSF independent variables (including PPVT-R, VMI, DAP PLS I through PLSV and MAS I through MAS III) as well as (2) only the 8 EPSF subtests as independent variables to predict Gates MacGinitie Vocabulary, Comprehension and Total Reading scores. Thus, two extra stepwise multiple regression analyses were done to further determine the contributing effects of various combinations of EPSF independent variables. Supplemental multiple regression #1 basically looked at the same independent variables were used in the research question two multiple except that the PLS Total and MAS Total were dropped as potential EPSF independent variables. Supplemental multiple regression analysis #2 basically looked at the predictive capabilities of the PLS and MAS subtests without the 5 EPSF screening tests.

A summary table of the various EPSF independent variables involved in the supplemental multiple regression analysis #1 and #2 a well as those in research questions one and two was generated (see Appendix I). It can be noted in Appendix I that the overall adjusted R squared or amount of variance accounted for by the EPSF variables changed approximately less than 1 % when the PLS Total and MAS Total when not included as predictors of Gates MacGinitie Vocabulary, Comprehension and Total Reading scores. Also, it was noted that the 8 EPSF subtests alone definitely resulted in significant decrease in the amount of variance in the Gates MacGinitie Vocabulary, Comprehension and Total reading scores predicted by the EPSF independent variables.

The noted significant stepwise multiple regression predictors (i. e. "significance in F change") comparisons between the EPSF 11 independent variable set (including PPVT-R, VMI, DAP, PLS I through V and MAS I through MAS III) and EPSF 13 independent variables set (including all 5 EPSF screening tests and the 8 total MAS and PLS subtests) showed that the significant predictors basically remained very similar between the two sets for predicting Gates MacGinitie Vocabulary, Comprehension and Total Reading score (see Appendix J). It can be noted from the various EPSF independent variables sets in Appendix J that basically 2 to 4 EPSF predictors in each set effectively predict Gates MacGinitie Vocabulary, Comprehension or Total Reading scores.

Appendix J shows that the deletion of PLS Total and MAS Total scores in the 11 EPSF independent variables set did not change the Gates MacGinitie Vocabulary or Total Reading score significant predictors or their order as compared to the 13 EPSF independent variables set. In effect, VMI, PPVT-R with PLS III remained (in that order) the best predictors of Gates MacGinitie Vocabulary scores and VMI, PPVT-R and PLS III remained (in that order) the best predictors of Gates MacGinitie Total Reading scores in both EPSF 13 and 11 independent variables sets. It is interesting to note that the DAP, MAS Total, and MAS I, II, and III are not seen as significant predictors in any EPSF variable set.

It was also noted in Appendix J that when comparing the 8 and 11 EPSF independent variables sets, the loss of PLS Total in the 8 EPSF variables set found the PLS III and PLS V contributing to the VMI and PPVT-R as significant predictors of Gates MacGinitie Comprehension. The loss of the PPVT-R and VMI as significant EPSF independent variable predictors in the 8 EPSF independent variables set (including only the 8 EPSF subsets) allowed the PLS subtests to increase in predictive potential. The PLS III and PLS V were noted significant predictors of Gates MacGinitie Vocabulary, Comprehension and Total Reading scores in the 8 EPSF independent variables set, if not other potential predictors beyond the EPSF subtests could be entered into the stepwise multiple regression. Still, overall less total variance (i.e., R Change) could be accounted for in all three Gates MacGinitie reading scores, if only EPSF subtests were used as the independent variables.

Research Question Three - The degree to which the kindergarten EPSF screening battery generated 7 developmental modality scores relate to and predict future, end of first grade, Gates MacGinitie Reading Tests achievement.

The third research question's statistical calculations were based on the sample of 630 subjects. Raw data was available for PLS and MAS calculations on only 373 of the total subjects. Thus only PLS subtests and MAS subtests comparisons to the 7 EPSF developmental scores could be done with the 373 subject sample. The obtained individual student EPSF diagnostic profile sheets (see Appendix A for sample profile) were available on all 630 subjects and were used to calculate the developmental modality scores used for research question three. The available EPSF tests and subtests as well as 7 developmental modalities data from the 373 subjects for research question three generated the following Pearson intercorrelational matrix (see Table 18).

Table 18
Pearson Intercorrelational Matrix of the 3 Gates MacGinitie Reading Scores, 5 EPSF Screening Battery Tests, 7 Developmental Modality Scores and 8 EPSF Subtests

	EPSF Developmental Modality Scores						
Variable	RL	EL	AU	VM	VD	FM	GM
GM Vocabulary	.2236**	.0791	.1066**	.1596**	.1198*	.1224*	.0587**
GM Comprehension	.3455**	.2654**	.2774**	.3052**	.3114**	.3181**	.0597
GM Total Reading	.3251**	.2416**	.2584**	.2807**	.3283**	.3327**	.0872
PPVT-R	.8351**	.5676**	.3239**	.1766**	.0757**	.0856	.0070
VMI	.0986	.0287	.1606**	.3005**	.6359**	.7628**	.0157
PLS Total	.4868**	.6652**	.5547**	.4938**	.2026**	.2313**	.0697
PLS 1	.4332**	.6088**	.2440**	.1460**	.0878	.0827	.0018
PLS 2	.3982**	.7040**	.2982**	.0908	.0069	.0165	.0054
PLS 3	.3435**	.3305**	.8575**	.2037**	.1507**	.2047**	.0739
PLS 4	.3931**	.5899**	.2692**	.1032**	.0608	.0661	-.0146
PLS 5	.1462**	.1684**	.1850**	.7399**	.2502**	.2706**	.1070*
MAS Total	.2081**	.2816**	.2212**	.0928	.0418	.0510	.3304**
MAS 1	.3545**	.3744**	.2758**	.0481	-.0195	-.0190	-.0556
MAS 2	.0650	.0923	.0688	.0717	-.0051	-.0077	.0053
MAS 3	-.0133	.0504	.1059*	.0599	.0880	.1054*	.4941**
DAP	.0766	.0272	.0581	.0331	.0163	.0013	.0312

GM = Gates MacGinitie ** = p<01 * = p<.05 N=373

Table 18 effectively shows the significant correlations in the 373 subject sample between the EPSF 7 developmental modalities scores, the 5 EPSF screening tests and 8 EPSF subtests. The most significant Pearson correlations above .60 (for the 373 sample) were in descending order: .85 for PLS III and AU, .83 for PPVT - R and RL, .76 for VMI and FM, .74 for PLS 5 and VMI, .70 for PLS II and EL, .67 for PLS Total and EL and .61 for PLS I and EL.

Table 19 shows (using the 630 subject sample) the specific Pearson correlational relationships between the EPSF 7 developmental modalites and the Gates MacGinitie Vocabulary, Comprehension and Total Reading scores.

Table 19
Pearson Intercorrelational Matrix of the 3 Gates MacGinitie Reading Scores and the EPSF 7 Developmental Modality Scores

Variable	1	2	3	4	5	6	7	8	9
2	.48**								
3	.53**	.93**							
4	.22**	.35**	.33**						
5	.08	.27**	.24**	.49**					
6	.11*	.28**	.26**	.32**	.32**				
7	.19**	.31**	.28**	.22**	.20**	.23**			
8	.12*	.31**	.33**	.18**	.11*	.18**	.40**		
9	.12*	.32**	.33**	.17**	.10*	.21**	.38**	.84**	
10	.06	.06	.09	-.04	.01	.08	.08	.17**	.05

1 = Gates MacGinitie Vocabulary, 2 = Gates MacGinitie Comprehension, 3 = Gates MacGinitie Total Reading, 4 = Receptive Language (RL), 5 = Expressive Language (EL), 6 = Auditory (AU), 7 = Visual Memory (VM), 8 = isual Discrimination (VD), 9 = Fine Motor (FM), 10 = Gross Motor (GM)

** = p<.01 * = p<.05
N = 630

The next level of data interpretation for research question three involved the use of canonical statistical analysis to create a synthetic composite linear dependent variable including the Gates MacGinitie Vocabulary, Comprehension and Total Reading scores as predicted by the derived composite linear independent variable comprised of the 7 EPSF modality scores. The SSPS statistical package yielded two canonical variants with only canonical variant root I being statistically significant with a Chi Square of 157. 169, df of 14 and p of .001 (see Table 20). Also, a preliminary check for non - linearity of the test data had been done during the initial data analysis phase.

Table 20
Canonical Results of the 7 EPSF Derived Modality Scores and the Gates MacGinitie Vocabulary and Comprehension Subtests Composite

Canonical Variant	Canonical Correlation	Canonical Correlation Squared	CHI Square	df	p
I	.46243	.21385	157.169	14	<.001
II	.10740	.01153	7.271	6	>.200

The canonical analysis results showed that canonical variant I was noted to exhibit 21.38% of the shared variance between the defined dependent and independent linear variable composites. Also, canonical variant I exhibited a .46 correlation between the dependent and independent linear variables composites. Canonical variant II was not found to be significant and thus only canonical root I was retained for further analysis. Obtained canonical variant structure coefficients or correlations (i. e., loadings) were calculated (see Table 21). Standardized variant canonical function coefficients (i. e., weights) are noted in Appendix K.

Table 21
Canonical Variant Structure Coefficients of the 7 EPSF Modality Scores and Gates MacGinitie Vocabulary and Comprehension Scores

Independent Variable		Dependent Variable	
RL	.73	Gates MacGinitie Vocabulary	.52
EL	.54	Gates MacGinitie Comprehension	.99
AU	.66		
VM	.60		
VD	.56		
FM	.56		
GM	.04		

Table 21 shows that 6 of the 7 EPSF independent variable canonical loadings were deemed significant (i. e., above a .30 level) with RL having the highest canonical correlational loading. Also, Table 21 noted that of the dependent variable loadings that Gates MacGinitie Comprehension had the most significant dependent variable correlation with the defined canonical variant with a noted level of .99.

Redundancy coefficients were computed for the obtained significant canonical variant I. The current canonical analysis data determined that given the dependent variable linear composite of Gates MacGinitie Vocabulary and Comprehension, that 32.35 % of the variance can be accounted for or explained by the linear composite independent variable (composed of the 7 EPSF developmental modalities). Also, given the EPSF 7 developmental modalities variable linear composite, that approximately 60. 42% of the variance can be accounted for or explained by the Gates MacGinitie Vocabulary and Comprehension dependent variable linear composite.

Follow-up stepwise multiple regression statistical analysis allowed for potential inclusion of the Gates Total Reading score as a dependent variable. The Gates MacGinitie Total Reading score, as noted earlier, was not used as part of the generated canonical dependent variable linear composite due to its being an additive function of the Gates MacGinitie Vocabulary and Comprehension subtests and thus, potentially distorting the linear canonical dependent variable composite.

The initial stepwise multiple regression analysis for research question three involved the prediction of Gates MacGinitie Vocabulary test scores (i.e., the dependent variable) by the 7 EPSF developmental modality scores (i.e., the independent variables). The only noted significant predictors of future Gates MacGinitie Vocabulary subtest scores were found to be RL and VM which were at the .01 and .05 levels of significance, respectively (see Table 22). The obtained Beta values for RL and VM were only .19 and .10, respectively. Still as compared to the other 5 EPSF developmental scores, RL and VM were significant predictors of future Gates MacGinitie Vocabulary achievement.

Table 22

Betas, b-Weights, Standard Errors (S.E.), t Ratios and p Values
for the Gates MacGinitie Vocabulary Scores as Predicted
by the 7 EPSF Derived Developmental Modality Scores

Variable	Beta(B)	b-Weight	S.E.	t Ratio	p Value
RL	.1908	8.2978	1.9628	4.227	.0000
VM	.1016	6.7781	2.8789	2.354	.0189
RM	.0482	3.2859	5.2212	.629	.5294
AU	.0479	2.2975	2.0558	1.118	.2642
EL	-.0363	-2.1398	2.6801	-.798	.4249
GM	.0305	.7903	1.0136	.780	.4359
VD	-.0027	-.2234	6.5924	-.034	.9730
Y intercept -7.5280					

df = 622

Table 23 shows that overall the 7 EPSF developmental modality scores only accounted for 6.6 % of the total variance in the Gates MacGinitie Vocabulary subtest scores. The adjusted R squared (i.e., shrinkage) determined that only a total of 5.5% of the variance in Gates MacGinitie Vocabulary scores could be accounted for by the 7 EPSF developmental scores if used with a different sample. Still of the 6.6% of total variance accounted for by the current sample 7 EPSF developmental scores, over 6% (specifically 6.04 %) of the total variance was accounted for by the combined RL and VM developmental modalities. Thus, the addition 5 EPSF developmental modality scores only added .57 of 1 % of the total variance accounted for by the entire 7 EPSF developmental modality scores. Therefore, the combined RL and VM developmental modalities accounted for approximately 92% of the total variance noted for in Gates MacGinitie Vocabulary achievement by using all 7

EPSF developmental modalities.

Table 23

Multiple Regression for the Gates MacGinitie
Vocabulary as predicted by the 7 EPSF Developmental Scores

Step	Variable Entered	R Squared	Increase R Square	Adjusted R squared	F Equation	F Equation Change	Sign. of F Change
1	RL	.0457	.0457	.0442	30.081	30.081	.000
2	VM	.0604	.0147	.0574	20.144	9.780	.002
3	FM	.0627	.0023	.0582	13.948	1.521	.218
4	AU	.0641	.0015	.0582	10.710	.999	.318
5	EL	.0651	.0010	.0576	8.693	.646	.422
6	GM	.0660	.0009	.0570	7.341	.611	.435
7	VD	.0660	.0000	.0555	6.282	.001	.973

N = 622

The second stepwise multiple regression analysis for research question three found that 3 EPSF developmental modalities were significant predictors at the .01 level for future Gates MacGinitie Comprehension subtest scores (see Table 24). These three significant independent variable predictors and their respective beta scores were RL (.2125), AU (.1667) and VM (.1223).

Table 24
Betas, b-Weights, Standard Errors (S.E.), t Ratios and p Values
for the Gates MacGinitie Comprehension Scores as Predicted
by the 7 EPSF Developmental Modality Scores

Variable	Beta(B)	b-Weight	S.E.	t Ratio	p Value
RL	.2125	4.5975	.8973	5.124	.0000
FM	.1013	3.4407	2.3869	1.442	.1499
AU	.1667	3.9803	.9398	4.235	.0000

Table 24 (continued)
Betas, b-Weights, Standard Errors (S.E.), t Ratios and p Values
for the Gates MacGinitie Comprehension Scores as Predicted
by the 7 EPSF Developmental Modality Scores

Variable	Beta(B)	b-Weight	S.E.	t Ratio	p Value
VM	.1223	4.0591	1.3161	3.084	.0021
EL	.0514	1.5086	1.2252	1.231	.2187
VD	.0633	2.6792	3.0137	.889	.3743
GM	.0062	.0792	.4634	.171	.8644
Y Intercept = - 11.0016					

df = 622

It was determined that overall the 7 EPSF developmental modality scores accounted for 21.2 % of the total variance in the Gates MacGinitie Comprehension subtest score (see Table 25). The adjusted R squared (i.e., shrinkage) found that 20.3 % of the total variance in an individual's Gates MacGinitie Comprehension subtest score (i.e., the dependent variable) could be accounted for by the 7 EPSF developmental modality scores (i.e., the independent variables) if used with a different sample.

Table 25
Multiple Regression Results for the Gates MacGinitie Comprehension
as Predicted by the 7 EPSF Developmental Scores

Step	Variable Entered	R Squared	Increase R Squared	Adjusted R squared	F Equation	F Equation Change	Sign. of F Change
1	RL	.1127	.1127	.1113	79.762	79.762	.000
2	FM	.1617	.0490	.1590	60.477	36.663	.000
3	AU	.1949	.0332	.1911	50.528	25.839	.000
4	VM	.2091	.0142	.2041	41.317	11.211	.001
5	EL	.2110	.0019	.2047	33.382	1.508	.220
6	VD	.2121	.0011	.2045	27.951	.838	.360
7	GM	.2121	.0000	.2033	23.925	.029	.864

N = 622

The combined RL, FM, VM and AU developmental modalities were noted to account for 20.91 % of the total (independent variables) variance in the Gates MacGinitie Comprehension subtest. Less than 1% of the total variance accounted for in the Gates MacGinitie Comprehension results was obtained from the remaining 3 EPSF developmental modalities. In effect, approximately 99% of the total variance accounted for in Gates MacGinitie Comprehension can be credited to the RL, FM, VM and AU developmental modalities. Thus, elimination of the EL, VD, and GM developmental modalites would not dramatically decrease the obtained Gates MacGinitie Comprehension subtest variance accounted for by the remaining RL, FM and AU modalities.

The final portion of research question three involved using stepwise multiple regression statistical analysis in the prediction of the Gates MacGinitie Total Reading score by the 7 EPSF developmental modalities. It was noted from Table 26 that RL, AU and VM developmental modalities were significant at the .01 level as individual predictors of Gates MacGinitie Total Reading score achievement with respective beta loading levels of .2056, .1444 and .1153. The other 4 EPSF developmental modalities were noted to not be significant individual predictors of Gates MacGinitie Total Reading scores. The GM developmental modality continued to be the least effective EPSF developmental modality predictive variable for Gates MacGinitie reading test achievement (as was also noted in the Vocabulary and Comprehension multiple regression analyses).

Table 26
Betas, b-Weights, Standard Errors (S.E.), t Ratios and p Values for the Gates MacGinitie Total Reading Scores as Predicted by the 7 EPSF Derived Developmental Modality Scores

Variable	Beta(B)	b-Weight	S.E.	t Ratio	p Value
RL	.2056	4.3885	.8983	4.886	.0000
FM	.0976	3.2676	2.3894	1.368	.1719
AU	.1444	3.4010	.9408	3.615	.0003
VM	.1153	3.7751	1.3175	2.865	.0043
EL	.0434	1.2564	1.2265	1.024	.3061

Table 26 (continued)
Betas, b-Weights, Standard Errors (S.E.), t Ratios and p Values for the Gates MacGinitie Total Reading Scores as Predicted by the 7 EPSF Derived Developmental Modality Scores

Variable	Beta(B)	b-Weight	S.E.	t Ratio	p Value
VD	.0684	2.8560	3.0169	.947	.3442
GM	.0207	.2625	.4639	.566	.5716
Y Intercept = - 8.5822					

df = 622

Table 27 shows that 18.84 % of the total variance in the Gates MacGinitie Total Reading score can be accounted for by the 7 EPSF developmental modality scores. The adjusted R square (i. e., shrinkage) shows that 17.92 % of the total variance in Gates MacGinitie Total Reading scores can be accounted for by the 7 EPSF developmental modality scores if used with a different population. Thus, the noted total variance shrinkage would be less than 5 % from the original population to an hypothetical different population.

Table 27
Multiple Regression Results for the Gates MacGinitie Total Reading as Predicted by the 7 EPSF Developmental Scores

Step	Variable Entered	R Squared	Increase R Squared	Adjusted R squared	F Equation	F Equation Change	Sign. of F Change
1	RL	.1000	.1000	.0985	69.761	69.761	.000
2	FM	.1470	.0471	.1443	54.047	34.602	.000
3	AU	.1725	.0254	.1685	43.488	19.227	.000
4	VM	.1853	.0128	.1800	35.527	9.810	.002
5	EL	.1866	.0014	.1801	28.631	1.037	.309
6	VD	.1880	.0014	.1801	24.034	1.039	.308
7	GM	.1884	.0004	.1792	20.627	.320	.572

N = 622

A total of 18. 53 % (of the 18.84% obtained total variance) in Gates MacGinitie Total Reading achievement was accounted for by the EPSF developmental modalities of RL, FM, AU and VM. Thus, the other 3 EPSF developmental modalities of EL, VD and GM only accounted for less than 1 % of total variance in Gate MacGinitie Total Reading. In effect, the EPSF developmental modalities of RL, FM, AU and VM accounted for 98.35% of the obtained total variance (from all 7 EPSF developmental modalities) in Gates MacGinitie Total Reading achievement.

Summary

Research Question One

Significant statistical results were found regarding each of the three study research questions. Research question one involved the degree to which the 5 basic EPSF screening tests were related to and predicted future, end of first grade, Gates MacGinitie reading achievement. Three significant Pearson correlations were found among the 5 EPSF basic tests including: PLS Total to PPVT-R with a r. of .58, PLS Total to VMI with a r. of .17 and MAS Total to PLS Total with a r. of .46. THE DAP test was not significantly correlated to any other EPSF screening battery test (see Appendix C).

The DAP was the only 1 of the 5 EPSF screening battery tests that had no significant Pearson correlations with any of the Gates MacGinitie reading scores. All other 4 EPSF screening battery tests were significantly correlated at the .01 or .05 level to all three Gates MacGinitie reading scores (see Appendix D). The PPVT-R, VMI and PLS Total had very similar Pearson correlations in the .30's with the Gates MacGinitie Comprehension and Total Reading scores. The MAS had significant correlations in the .10's with all Gates MacGinitie reading scores.

Research question one canonical analysis found one significant canonical root with a .48 correlation of the 5 EPSF screening tests linear composite to the Gates MacGinitie Vocabulary and Comprehension linear composite. A shared variance of 23.39% between the dependent and independent variable canonical composites was noted. All the EPSF screening tests except DAP were noted significant canonical predictors or loadings for the Gates MacGinitie Vocabulary and Comprehension Composite. The PPVT-R, PLS Total and VMI tests had similar canonical correlations of .71, .73 and .69, respectively, with MAS (Total) noted to be significant at the .31correlational level.

Other noted research question one significant canonical factors included the dramatic .99 factor loading and .99 canonical weight for the Gates MacGinitie Comprehension subtest (see Table 5 and Appendix E). The redundancy coefficients showed the major impact of Gates MacGinitie Vocabulary and Comprehension linear composite with 60.28 % of the variance in the 5 EPSF screening test linear composite could be explained by the Gates MacGinitie reading tests composite (versus 32.92% of the variance in the Gates MacGinitie reading tests composite could be explained by the linear composite of the 5 EPSF screening tests).

Pearson correlations for research question one stepwise multiple regression analysis found the most significant intercorrelation in both the 373 and 633 sample sets of the Gates MacGinitie 3 reading scores to be .93 between the Gates MacGinitie Comprehension and Gates MacGinitie Total Reading score (see Appendix F). This again showing the impact of the Gates MacGinitie Comprehension subtest in respect to the overall Gates MacGinitie reading test. Overall the amount of total shared variance (i.e., R Squared) accounted for by all 5 EPSF with the Gates MacGinitie Vocabulary, Comprehension and Total Reading scores were 5.5%, 23.38% and 22.46%, respectively. Thus, the combined 5 EPSF screening battery tests were more significant predictors of Gates MacGinitie Comprehension and Total Reading scores than Gates MacGinitie Vocabulary.

Question one stepwise multiple regression analysis found that deletion of the DAP and MAS as overall EPSF screening test predictors of Gates MacGinitie Vocabulary, Comprehension and Total Reading scores would not significantly effect the overall prediction of Gates scores by the EPSF screening battery. In fact, 92%, 99% and 99% of the total variance accounted for in Gates MacGinitie Vocabulary, Comprehension and Total Reading scores, respectively, by the 5 EPSF screening battery tests was from the combined PPVT-R, VMI and PLS Total scores. Also, the shrinkage (i.e, adjusted R squared) of the predictive variance was minimal for the PPVT-R, VMI and PLS Total scores with only a 5% loss in total predicted variance for the Gates MacGinitie Comprehension and Total Reading scores, if used with a different population.

The combined EPSF screening battery significant predictors for Gates MacGinitie Vocabulary, Comprehension and Total reading scores, respectively were PPVT-R and VMI; PLS Total, VMI and PPVT-R and VMI, PPVT-R and PLS Total. Overall, the MAS Total and DAP could be effectively dropped from the EPSF screening battery with minimal loss in predictive capabilities for Gates MacGinitie Vocabulary, Comprehension or Total Reading scores (see Tables 5, 7 and 9).

Research Question Two

Research question two involved determining the extent to which the 8 EPSF screening battery subtests were related to and contributed to the basic 5 EPSF screening battery's prediction of end of first grade Gates MacGinitie reading achievement. Pearson intercorrelations found that the 5 PLS subtests are correlated in the .60 range to PLS Total and the MAS I, II and III subtests correlated . 74, .16 and .74 with MAS Total, respectively (Table 1). Overall, all PLS and MAS subtests were correlated significantly at the .01 level with their own respective test total (be it MAS Total or PLS Total).

Pearson intercorrelations also found that MAS II and MAS III were not significantly correlated with any of the three Gates MacGinitie reading scores (see Appendix G). In fact, MAS II was negatively correlated to Gates MacGinitie

Comprehension and Total Reading scores. MAS I was significantly correlated at the .01 level to all three Gates MacGinitie reading scores. In contrast, PLS III and V were at the .01 significance level for Pearson correlations with all three Gates MacGinitie reading scores. PLS I, II and IV were at the .01 significance level when correlated with Gates MacGinitie Comprehension and Total Reading scores.

Research Question two canonical analysis found (when the combined 5 EPSF screening tests and 8 EPSF subtests were used as the linear independent composite) that the correlation between the synthetic EPSF variable and the combined Gates MacGinitie Vocabulary and Comprehension subtests (dependent variable) was .51 (as compared to a .48 when the 5 EPSF screening tests were the independent canonical variable). The proportion of shared variance (i.e, canonical correlation squared) between the EPSF linear composite and Gates reading composite increased from 23.38% (when the 5 EPSF screening tests were used alone as a composite linear predictor) to 25.85% when the 8 EPSF subtests were added to the canonical independent variable composite. Thus the addition of the 8 EPSF subtests to the 5 EPSF screening battery tests (as a independent canonical linear composite) only increased the total shared variance 2.53% or approximately a 10% increase in total variance accounted for by adding (to the research question one canonical analysis) the 8 EPSF subtest scores in research question two.

The canonical redundancy coefficients found only an increase from 60.26% to 60.82% of the variance explained in the EPSF independent linear composite by the Gates MacGinitie Reading tests composite when the 8 EPSF subtests were added to the original 5 EPSF screening tests. Interestingly enough, the redundancy index for the amount of variance explained by the EPSF independent variable linear composite given the Gates reading tests composite dropped from 32.92% (when the 5 EPSF screening tests were the independent composite) to 20.35% (when the 8 EPSF subtest scores were added to the 5 EPSF tests to form the amended research question two independent variable linear composite).

The significant independent variable canonical loadings or predictors for the combined Gates MacGinitie Vocabulary and Comprehension linear dependent variable composite included all 5 EPSF PLS subtests and all EPSF screening tests except DAP. Thus, the MAS I, II and III subtests were not significant canonical predictors of the combined Gates MacGinitie Vocabulary and Comprehension reading tests linear composite.

The most significant canonical individual EPSF predictors of the Gates reading composite achievement were PLS Total, PPVT-R, VMI, PLS III with correlations of .69, .68, .65 and .61 followed by PLS I, PLS V, PLS IV, PLS III and MAS Total with correlations of .45, .44, .40, .34 and .30, respectively. Thus, the addition of the 8 EPSF subtest scores (to the 5 EPSF screening tests) did increase the number of significant canonical independent variable loadings or predictors of combined Gates MacGinitie Vocabulary and Comprehension scores. Still, the addition of the 8 EPSF subtests to the 5 EPSF subtests to form the canonical independent variable (1) only added a 2.53% increase (from the 5 EPSF tests alone) in total variance accounted for by the created (13 variable) EPSF tests and subtests linear composite and (2) caused a drop (from 32.92% using the 5 EPSF tests alone) to 20.32% in the total variance explained by the independent variable composite, given the Gates reading test dependent composite.

Research question two multiple regression analysis found that addition of the 8 EPSF subtests to the 5 EPSF tests did increase the overall total variance accounted for in the Gates Vocabulary, Comprehension and Total Reading scores by the EPSF independent variables (see Table 22). Supplemental multiple regression data analysis was done using (1) the 8 EPSF subtests only and (2) the 8 EPSF subtests with PPVT-R, VMI and DAP (without the PLS Total and MAS Total scores) to help determine the magnitude of effects of the EPSF subtests in predicting future Gates MacGinitie reading achievement.

Overall, the addition of the 8 EPSF subtest scores to the original 5 EPSF test scores did raise the total amount of variance predicted for the Gates MacGinitie Vocabulary scores from 5.5 % (when considering the 5 EPSF screening tests alone) to 8.0 %. Thus ,the amount of total variance gained by the addition of the 8 EPSF subtest scores as predictors of Gates MacGinitie Vocabulary achievement did not dramatically change from the 5 EPSF screening tests alone as predictors of future Gates Vocabulary achievement. In fact, when considering the number of significant combined EPSF predictors of future Gates Vocabulary achievement, the amount of total variance accounted for in Gates Vocabulary scores was noted to be 5.08% by the PPVT-R and VMI tests alone (see Table 5) versus addition of the significant individual EPSF subtest scores (in any combination or sequence) to the PPVT-R and VMI only produced an overall significant independent variable prediction level of 5.43% for Gates Vocabulary scores (see Appendix J).

The supplementary stepwise multiple regression analysis using (1) the 8 EPSF subtests only and (2) the 8 EPSF subtests and EPSF screening tests (without the MAS Total and PLS Total) found that the total variance in Gates MacGinitie Vocabulary scores predicted by the created 8 and 11 independent variable sets (see Appendix G) to be 4.42% and 7.84%, respectively. The PLS V and PLS II subtests were noted to become significant predictors of Gates MacGinitie Vocabulary scores when the 8 EPSF subtests tests were added to the 5 EPSF tests as potential EPSF predictors (see Appendix H). The PPVT-R, PLS V and PLS II were the significant predictors when the 8 EPSF subtests were added to the 5 EPSF screening tests as potential predictors of Gates MacGinitie Vocabulary. A reduced prediction model for Gates MacGinitie Vocabulary including PPVT-R and VMI or PPVT-R, PLS V and PLS II would predict 5.08 % and 5.43%, respectively, of the total variance in Gates MacGinitie Vocabulary scores.

The initial significant predictors (found in the 5 EPSF screening tests) including PLS Total, VMI and PPVT-R were noted to account for (in varied order) a total of 23.18 % and 22.35 % in Gates MacGinitie Comprehension and

Total Reading scores, respectively (see Tables 7 and 9). The inclusion of the 8 EPSF subtests scores (in various EPSF tests and or subtests combinations) to predict future Gates MacGinitie Comprehension and Total Reading Scores did not increase the amount of significant total variance predicted by the PLS Total, PPVT-R and VMI alone nor change their significance as the best combined predictors of Gates MacGinitie Comprehension and Total Reading scores (see Tables 7 and 9, as well as Appendix G and Appendix H). The MAS Total , three MAS subtests and DAP showed no significant contribution to the prediction of the Gates MacGinitie Vocabulary, Comprehension or Total Reading scores.

The Pearson correlations between the PLS and MAS subtests found the highest significance correlations at the .05 level with the Gates MacGinitie Vocabulary by the PLS V, PLS III and MAS I with correlations of .15, .13 and .12, respectively with a range of .01 to .15 including all PLS and MAS subtests (see Table 1). In comparison, the PPVT-R, PLS Total, VMI and MAS Total had significant correlations with Gates Vocabulary of .19, .16, .13 and .12, respectively. Thus, overall while statistically significant, the Pearson correlations between the EPSF tests and subtests with Gates MacGinitie Vocabulary were at best below .20.

Interestingly enough, when the PLS and MAS subtests were added potential significant EPSF predictors in research question two, the same three EPSF basic screening tests of PLS Total, VMI and PPVT-R remained the best combined predictors for total variance in the Gates MacGinitie Comprehension (see Appendix H). If PLS Total was dropped from the EPSF independent variables, then PLS III and PLS V became significant predictors of Gates MacGinitie Comprehension scores. Still, the PLS Total, VMI and PPVT-R together produced a combined predictor variance for Gates MacGinitie Comprehension of 23.18% versus 22.95% and 12.96% when the EPSF tests/subtests set without PLS Total and MAS Total scores and (2) the 8 EPSF subtests alone, respectively, were considered (see Table 7 and Appendix H).

The addition of the EPSF 8 subtest scores (in the calculation of the total variance predicted by the EPSF combined tests and subtests) allowed the PLS III to replace the PLS Total as the third most significant predictor of Gates MacGinitie Total Reading scores when combined with the VMI and PPVT-R. Thus the PLS III shows its potential as a useful significant predictor of Gates MacGinitie Total Reading score but still it only adds less than 2% of the total variance accounted for by the more powerful predictors of VMI and PPVT-R.

Research Question Three

The third research question involved the extent to which the EPSF 7 developmental modality scores were related to and predicted future, end of first grade, Gates MacGinitie reading achievement. Pearson correlations (using both samples) showed that 6 of the 7 EPSF developmental modality scores (excluding GM) were significantly correlated at the .05 level to Gates MacGinitie Comprehension and Total Reading scores (see Tables 18 and 19).

The Gates MacGinitie Vocabulary subtest score was significantly correlated at the .01 level with 4 EPSF modality scores including RL, VM, FM and AU. The VD modality was correlated at the .05 level to the Gates MacGinitie Vocabulary subtest scores. The GM and EL modalities were not significantly correlated to the Gates MacGinitie Vocabulary scores. The RL, FM, AU and VD modalities were significantly correlated with all three Gates MacGinitie reading scores. The GM modality score was not significantly correlated with any of the three Gates MacGinitie reading achievement scores.

The most significant Pearson correlations (i.e. above a correlation of .60) among the EPSF developmental modalities and EPSF screening tests and subtests were in descending order: .85 for PLS III and AU; .83 for PPVT-R and RL; .76 for VMI and FM; .74 for PLS V and VMI; .70 for PLS II and EL; .67 for PLS Total and EL; and .61 for PLS I and EL (see Table 18). It is interesting to note that the seven aforementioned significantly correlated EPSF subtest and EPSF developmental modality combinations are all noted in the EPSF program developers' defined appropriate EPSF developmental modalities and their defined specific EPSF subtests (see Dissertation Chapter III, page 54).

Research Question three canonical analysis found that a .46 canonical correlation existed between the defined canonical linear dependent variable set (composed of the 7 EPSF modality scores) and the canonical linear dependent variable set of the Gates MacGinitie Vocabulary and Comprehension subtest scores. In comparison, canonical correlations of .48 and .51 were found when the 5 EPSF screening test scores and combined 13 EPSF tests and subtests' scores were used, respectively, as canonical dependent linear variable sets.

A shared variance (i.e., canonical correlation squared) of 21.38% was found between the defined EPSF 7 modality scores dependent linear composite and the defined independent variable linear composite of the Gates MacGinitie Vocabulary and Comprehension subtest scores. In contrast, the shared variance between the Gates MacGinitie dependent variable linear composite and the dependent variable linear composites of the EPSF 5 screening tests or combined 13 EPSF tests and subtests were 23.38% and 25.81%, respectively. Thus, the use of the EPSF screening tests alone would not loss appreciable total variance accounted for in predicting combined Gates MacGinitie Vocabulary and Comprehension subtest scores (as compared to independent canonical variable linear composites formed by the 7 EPSF developmental scores or 13 total EPSF tests and subtests' scores).

The defined significant independent variable canonical root structure coefficients or loadings were (in descending order of significance) RL, AU, VM, VD, FM and EL with GM not deemed a significant loading (see Table 21). The Gates MacGinitie Comprehension subtest was again (as noted in research questions one and two) a very significant dependent variable canonical loading with a correlation of .99. The RL and AU modalities were the most significantly weighted canonical independent variable functions with .47 and .36 correlations, respectively.

The calculated canonical redundancy coefficients found that given the

EPSF 7 developmental scores linear composite, that 60.42% of its variance can be accounted for by the Gates MacGinitie Vocabulary and Comprehension subtests linear dependent variable composite (as compared to 60.28% and 60.82% accounted for in the EPSF 5 screening tests and 13 combined EPSF tests and subtests, respectively). The 32.35% amount of variance accounted for in the Gates MacGinitie Vocabulary and Comprehension linear composite by the defined independent variable composite of the 7 EPSF modality scores is very similar to the 32.92% of dependent variable linear composite variance accounted for by the 5 EPSF screening tests.

Research question three stepwise multiple regression analysis found that overall 6.6% of the total variance in the Gates MacGinitie Vocabulary score can be accounted for by the combined 7 EPSF developmental modality scores (see Tables 22). In contrast, the overall total variance in Gates MacGinitie Vocabulary scores was better accounted for by the independent variable sets of (1) the 13 total EPSF tests and subtests and (2) the 11 EPSF variable set composed of the 8 EPSF subtests and 3 EPSF tests (excluding PLS Total and MAS Total) with 8.0% and 7.84% , respectively, of the total variance (see Appendix G).

The overall total variance accounted for by the total 7 EPSF modality scores in the Gates MacGinitie Comprehension and Total Reading scores were found to be 21.21% and 18.84%, respectively. Previous research question one and two calculations found that both the EPSF 5 screening battery tests and combined 13 EPSF variable set (of all EPSF subtests and tests) better predicted overall Gates MacGinitie Comprehension and Total Reading scores versus Gates MacGinitie Vocabulary scores than the EPSF modality scores (see Appendix H).

The noted best single independent variable predictors (i.e., beta loadings) for Gates MacGinitie Comprehension and Total Reading scores were all at the .05 level of significance and identical modalities in the same identical order of RL, AU and VM. The most noted single independent variable predictor (i.e.., beta loadings) for Gates MacGinitie Vocabulary were in order RL and VM. Overall, RL was the best EPSF modality single predictor of Gates MacGinitie Vocabulary, Comprehension and Total Reading scores (see Tables 22, 24 and 26).

The combined EPSF modalities that accounted for the most variance in the Gates MacGinitie Vocabulary subtest were RL and VM with 6.04 % of the total variance. The most significant combined EPSF modality scores that accounted for the most variance in Gates MacGinitie Comprehension and Total Reading scores were identical and in the descending order of RL, FM, AU and VM (see Tables 23, 25 and 27). The initial RL modality accounted for at least one half of the total significance variance accounted for by the four significant modality predictors of Gates MacGinitie Comprehension and Total Reading scores. Overall 20.91 % and 18.53% of the total variance accounted for in Gates MacGinitie Comprehension and Total Reading scores were contributed to the RL, FM, AU and VM significant independent variable predictors. The EPSF modalities of GM, EL and VD were noted to not significantly contribute as individual predictors of any of the defined three Gates MacGinitie reading scores.

Chapter V

Discussion, Summary and Conclusions

This chapter summarizes the purpose, methodology and findings of this study. Conclusions based upon the study findings and implications will be discussed. Recommendations are made regarding further research and current study concerns.

Purpose of Study

This study was undertaken to determine the predictive validity capabilities of the Early Prevention of School Failure (EPSF) screening battery, its components and derived developmental modality scores in the prediction of future Gates MacGinitie reading achievement. The EPSF screening battery has been used since the early 1970's as an integral part of the EPSF screening and intervention program. The overall EPSF program is a nationally validated preschool developmental screening program designed to prevent school failure through early identification of 4 to 6 year children's developmental skills and learning styles (Werner, 1990).

The EPSF designated "curriculum design" and "classroom management" program components are used in a developmental deficit remedial program (used typically in the kindergarten setting) to help identified children at risk for school failure increase their developmental proficiencies in seven EPSF developmental areas through 52 EPSF developed reading and writing learning tasks. Specifically, Werner (1990) stated " the EPSF nationally validated program identifies each child's developmental levels and learning styles for the teacher initiated child centered developmental learning experiences in the classroom", (p. 1).

Despite the widespread use of the EPSF intervention program in over 2,000 school districts located in 48 state and in five foreign countries (Driscoll, 1992,) with over 500,000 young children (Bryant, 1991), only 5 independent studies were noted in the literature involving predictive validity of the EPSF screening battery (Agostin, 1993; Bryant, 1991; McConnell, 1986; Terbush, 1990 and Roth,et al. 1993). Numerous researchers have commented on the need for evidence of predictive validity in kindergarten screening measures (Lehr, Ysseldyke & Thurlow, 1987; Lichenstein & Ireton, 1984; Lindsay & Wedell, 1982; McConnell, 1986; and Meisels, 1985). Meisels, Wiske & Tivnan (1984) stated "most developmental screening instruments provide extremely limited validity information and very few describe the relationship between screening data and later school achievement", (p, 25).

The current study examined the predictive validity capabilities of the EPSF 5 screening battery tests, its 8 subtests and the derived 7 EPSF developmental modality scores as related to and predictors of future Gates MacGinitie Reading Test achievement. The Gates MacGinitie is a well known reading achievement test developed in 1926 with revisions in 1976 and 1989. The Gates MacGinitie has been called a "prototype of the contemporary standardized reading test", (Calfee, 1985, p. 593). The Gates MacGinitie is reported in the literature as an effective general screening of reading abilities (Calfee, 1985; Cooter & Curry, 1989; Lindquist, 1982; Graham, 1990 and MacGinitie & MacGinitie 1989a). The Gates MacGinitie is used in school systems throughout the United States as a general reading screening for grades kindergarten through twelve.

The present study involved the following three research questions:

1. What is the degree to which the kindergarten age administered EPSF screening battery 5 tests are related to and predict future, end of first grade, Gates MacGinitie reading achievement?

2. What is the extent to which the kindergarten age administered EPSF screening battery 8 subtests are related to and contribute to the EPSF basic screening battery 5 tests' prediction of future, end of first grade, Gates MacGinitie reading achievement?
3. What is the degree to which the kindergarten age administered EPSF screening battery generated 7 individual developmental modality scores are related to and predict future, end of first grade, Gates MacGinitie reading achievement?

Methodology

The subjects in this current study were two samples of 373 and 630 children entering kindergarten in six different schools in a midwestern community of approximately 35,000 residents. The 373 subject sample is a subset of the overall 630 subjects group. Both samples were similar in age and sex ratio with the average age of the subjects in the 373 and 630 sample groups being 66.62 months and 66.68 months ,respectively, at time of screening. Males comprising 54% of the 373 sample set and 56% of the 630 group sample.

The approximately 325 to 400 potentially entering kindergarten children had been routinely EPSF screened as a part of the public school enrollment process in the community used in the present study. Entering kindergarten children EPSF screening results were gathered for the 1990 - 1991, 1991 - 1992 and 1992 - 1993 school years with follow-up April, 1992, 1993 and 1994 first grade Gates MacGinitie achievement scores gathered. Thus, an entering kindergarten student's EPSF screening battery results and generated EPSF developmental profile were used to predict their eventual end of first grade Gates MacGinitie reading achievement scores. Approximately one half of the initial kindergarten EPSF tested children were not available for follow -up first grade Gates MacGinitie testing due to noted student attrition caused, in part, to the highly mobile school age population resulting largely from the annual influx and departure to the community of the local university students and their families.

The EPSF computer program scored developmental profile sheets for each student (see Appendix A) were available for all 630 subjects with individual EPSF raw data found for 373 of those subjects. The total age and raw score developmental age equivalences for the PPVT-R, VMI and EPSF.DAP were collected from the 630 EPSF developmental profiles with total raw scores calculated for the EPSF.PLS and EPSF.MAS on the available 373 subjects screening sheets. The available EPSF developmental profiles did not show the EPSF.MAS and EPSF.PLS raw scores. The 630 EPSF developmental profile sheets contained each subject's generated 7 EPSF developmental modality scores. These seven scores were each numerically ranked from 1 to 5 based on their individual corresponding modality strength or need level.

Statistical analysis of the predictive capabilities of the 5 EPSF tests, 8 subtests and 7 developmental modalities were eventually used to predict future, end of first grade, Gates MacGinitie Total Reading, Vocabulary and Comprehension scores. Initial data analysis on each of the three research questions included generation of appropriate Pearson intercorrelational matrices followed by canonical analysis of the specific composite EPSF independent variable sets in each research question in relationship to the Gates MacGinitie Vocabulary and Comprehension scores composite. Initial canonical analysis was used to eliminate any potential loss of valuable information from the independent or dependent variables caused by data analysis of the variables as separate entities.

Secondary stepwise multiple regression analysis was preformed in each of the three research questions. The use of stepwise multiple regression allowed the inclusion of the Gates MacGinitie Total Reading score in the data analysis. This was not possible in the canonical analysis due to the fact that the Gates MacGinitie Total Reading scores is an additive function of the other two

Gates MacGinitie scores (thus, if used, potentially unfairly weighting the canonical dependent variable linear composite). The stepwise multiple regression analysis on all three research questions looked at the specified EPSF independent variables in each research question as predictors of future Gates MacGinitie Total Reading, Vocabulary and Comprehension scores as separate dependent variables.

Discussion

Research Question One

The initial research question dealt with the degree to which the kindergarten administered EPSF 5 basic screening tests were related to and predicted future, end of first grade, Gates MacGinitie reading achievement. Initial Pearson intercorrelations found the PLS Total test to be significantly correlated with three other EPSF tests including PPVT-R, VMI and MAS Total. The DAP was not correlated with any of the other 4 EPSF screening tests nor was it significantly correlated with any Gates MacGinitie reading score. The other 4 EPSF screening tests were all significantly correlated to all three Gates MacGinitie reading scores. Overall, the PPVT-R, VMI and PLS Total were the highest Pearson correlated EPSF screening tests to all three Gates MacGinitie reading scores with MAS and DAP having lower Gates MacGinitie intercorrelations to all three Gates MacGinitie reading scores.

Canonical analysis found a correlation of .48 between the combined 5 EPSF variables and the combined Gates Vocabulary and Comprehension variable composite with 23.39% shared variance between the two sets. All of the 5 EPSF screening test scores (except DAP) were significant individual canonical predictors of the Gates MacGinitie Vocabulary and Comprehension linear composite with PPVT-R, PLS Total and VMI again noted to be closely related variables with loading ranging from .69 to .71 and Pearson correlations with each other in the .30s. MAS Total was the lowest of the 4 EPSF significant individual predictors of the Gates MacGinitie Vocabulary and Comprehension linear composite with a loading of .31.

The influence of the Gates MacGinitie Comprehension subtest as the major subtest component in Gates MacGinitie reading achievement was readily noted by its consistently high canonical loading of .99 for the Gates MacGinitie Vocabulary and Comprehension canonical linear composite. The Gates MacGinitie Comprehension subtest was also noted to have a .93 Pearson correlation with the Gates MacGinitie Total Reading score.

The major impact in each of the three research questions by the Gates MacGinitie Vocabulary and Comprehension canonical linear composite as predictors of the defined EPSF variable sets was readily noted. It was noted that 60.26% of the variance in the 5 EPSF screening tests was accounted for or explained by the Gates MacGinitie Vocabulary and Comprehension canonical composite.

The three research questions found that in the various defined EPSF variable sets could only at best account for or explain less that 33% of the variance in the Gates MacGinitie Vocabulary and Comprehension composite (versus 60.82% using the combined EPSF tests and subtests and 60.42% using the 7 EPSF modality scores). Research question one found 32.92% of Gates MacGinitie Vocabulary and Comprehension composite variance could be predicted by the 5 EPSF screening battery tests. Overall, the Gates MacGinitie Vocabulary and Comprehension composite was almost twice as effective in predicting EPSF screening battery test scores than vice versa.

It was noted in stepwise multiple regression analysis that 5.5%, 23.38% and 22.46% of the total variance in Gates MacGinitie Vocabulary, Comprehension and Total Reading scores, respectively, could be accounted for by the combined 5 EPSF tests. The EPSF tests of PPVT-R, VMI and PLS Total accounted for 92%, 99% and 99% of the total variance accounted for by the entire 5 EPSF screening tests as predictors of Gates MacGinitie Vocabulary, Comprehension and Total Reading scores, respectively. McConnell (1986) and Terbush (1990) noted in their EPSF studies that PPVT-R, VMI and PLS Total (in identical order) were the best predictors of Metropolitan Reading Test and IOWA Tests of Basic Skills (Reading) achievement , respectively.

The deletion of the MAS Total and DAP test scores from the 5 EPSF screening test battery could be effectively done without significant loss of predictive capabilities of future Gates MacGinitie Vocabulary and Comprehension in combination or all three Gates MacGinitie reading individual achievement scores. Overall the EPSF screening battery was much more effective in predicting Gates MacGinitie Comprehension and Total Reading (versus Vocabulary) scores.

The significant combined EPSF screening battery predictors of Gates MacGinitie Vocabulary scores were PPVT-R followed by VMI. PPVT-R has traditionally been noted in the literature as having a "median correlation with other vocabulary tests of .71" (McConnell, 1986, p. 44). Current study data found the PPVT-R to have a correlation of .19 and the highest Pearson correlation of the 5 EPSF screening tests to the Gates MacGinitie Vocabulary scores. VMI was the highest individual predictor of Gates MacGinitie Comprehension and Total Reading scores and just below PPVT-R as the most significant individual predictor of Gates MacGinitie Vocabulary. Wallbrown, et al. (1975) found that the Gates MacGinitie Comprehension score required a broader range of skills (especially including the visual motor area) than the Gates MacGinitie Vocabulary scores. If true, then the predictive impact of the VMI would be more readily noted in the Gates Comprehension and Total Reading (versus Vocabulary scores).

The VMI remained the best individual predictor (as noted by factor loadings) of Gates MacGinitie Comprehension and Total Reading scores followed by the PPVT-R and PLS Total. The stepwise multiple regression combined EPSF screening battery best predictors of Gates MacGinitie Comprehension and Total Reading scores were PLS Total followed by VMI and

PPVT-R. The PLS Total was the best of the combined 5 EPSF screening tests to predict Gates MacGinitie Comprehension but dropped to third place (behind VMI and PPVT-R) as the combined best EPSF predictor of Gates Total Reading score. Overall, the MAS Total and DAP were not significant individual or combined EPSF screening test predictors of any of the three Gates MacGinitie reading scores. Thus, the MAS Total and DAP could be effectively dropped from the 5 EPSF screening test battery as Gates MacGinitie reading score predictors without significant loss of predictive capacity.

Research Question Two

Research question two involved the addition of the 8 EPSF subtests (including PLS I through PLS V and MAS I through MAS III) to the basic 5 EPSF screening tests to determine how they related to or predicted future Gates MacGinitie reading achievement. Pearson correlations found all PLS and MAS subtests were significantly correlated with their respective PLS Total and MAS Total scores. Pearson correlations with the Gates MacGinitie reading scores found PLS III, PLS V and MAS I significantly correlated at the .01 level with all three Gates MacGinitie reading scores. PLS I, PLS II and PLS IV were significantly correlated with Gates MacGinitie Comprehension and Total Reading. MAS II and MAS III had no significant Pearson correlations with any of the three Gates MacGinitie reading scores. In fact, MAS II was the only EPSF test or subtest to be negatively correlated with Gates MacGinitie Comprehension and Total Reading.

The combined 8 EPSF subtests and 5 EPSF tests had a correlation of .51 with the Gates MacGinitie Vocabulary and Comprehension subtests canonical composite and a 25.85% shared variance between the two composite variable sets. Thus, only a 1.47% increase in shared variance (with the Gates MacGinitie Vocabulary and Comprehension composite) was noted when the 8 EPSF subtest scores were added to the 5 EPSF screening test scores to form the enlarged canonical independent variable composite. The addition of the 8 EPSF subtest scores to the 5 EPSF test scores actually caused a drop in amount of variance from 30.92% (using only the 5 EPSF test scores) to 20.32% explained in the combined Gates MacGinitie Vocabulary and Comprehension score composite by the enlarged EPSF combined tests and subtests composite. This drop in shared variance explained could be due, in part, to (1) the effects of the higher standard errors consistently found in the PLS and MAS subtests (as compared to the PLS Total and MAS Total scores) and (2) the negative effects of the addition of the non significant three MAS subtests to the other ten EPSF tests and subtests scores to form the canonical independent linear composite.

The most significant individual EPSF tests and subtests predictors (i.e.. factor loadings) for the Gates MacGinitie Vocabulary and Comprehension canonical linear dependent variable composite continued to be (the same as in research question one when only the 5 EPSF test scores were used) PPVT-R, VMI and PLS Total with .65 to .69 correlations and PLS III was significant at the .61 correlational level. PLS III was noted to have the highest of the PLS subtest Pearson correlations with Gates MacGinitie Comprehension and only PLS V having a higher PLS subtest Pearson correlation (than PLS III) with Gates MacGinitie Vocabulary.

The only noted significant EPSF individual predictor (of the possible 13 EPSF tests and subtests) for Gates MacGinitie Vocabulary was PPVT-R. PPVT-R was joined by PLS V and PLS II as the most significant combined EPSF test or subtest predictors of Gates Vocabulary. PLS V and PLS II are described by Weiner (1992b) as "Vocabulary" and "Discriminative Visual - Auditory Memory" subtests, respectively. PLS V was noted as the most highly Pearson correlated PLS subtest to Gates MacGinitie Vocabulary (see Table 1). Interestingly enough, PLS V, PLS III and PLS II were the three highest of five factors extracted from the children's responses used in the initial PLS Total factor analysis (Werner, 1992c).

Question two stepwise multiple regression found the total variance for

Gates MacGinitie Comprehension scores predicted by the 13 EPSF variable set of 5 tests and 8 subtests to be 23.13% (as compared to 23.38% total Gates MacGinitie Comprehension score variance predicted using the 5 EPSF subtests alone). Thus, the addition of the 8 EPSF subtests did not significantly change the amount of accounted variance predicted in Gates MacGinitie Comprehension by the 5 EPSF test scores.

The individual significant EPSF predictors of Gates MacGinitie Comprehension included PPVT-R and VMI at the .01 level of significance and PLS Total and all 5 PLS subtests (significant at the .05 level) . Thus, the 5 PLS subtests became significant individual predictors of Gates MacGinitie Comprehension when added to the 5 EPSF basic test scores. The best combined EPSF tests and subtests predictors of Gates MacGinitie Comprehension (as in research question one) are PLS Total, VMI and PPVT-R. No other complied supplemental EPSF subtest and /or subtests and tests variable combination better predicted Gates Comprehension then the PLS Total, VMI and PPVT-R.

PLS Total was the highest combined predictor of Gates MacGinitie Comprehension with 12% of the total 23.13% variance accounted for by the potential 5 EPSF tests and 8 EPSF subtests. PLS Total was noted to be the highest Pearson correlated EPSF test or subtest to the Gates MacGinitie Comprehension subtest. The MAS Total, all three MAS subtests and the DAP were not significant combined or individual predictors of Gates MacGinitie Comprehension scores.

A total of 24.43% of the total variance in the Gates MacGinitie Total Reading score was predicted by the combined EPSF tests and subtests with VMI, PPVT-R and PLS III accounted for 22.35% of the 24.43% variance noted. The addition of the 8 EPSF subtests did not dramatically increase the amount of total variance predicted in Gates MacGinitie Total Reading scores from the 22.46% variance originally accounted for by the 5 EPSF battery tests. In fact, the significant combined EPSF tests predictors (noted in research question one) of

VMI, PPVT-R and PLS Total by themselves accounted for 22.31% of the variance predicting in Gates MacGinitie Total Reading scores.

The increased significance of PLS III as a combined EPSF tests and subtests predictor (in research question two) of Gates MacGinitie Total Reading can be definitely noted in the fact the PLS III replaced PLS Total (in research question one) as the third most significant combined predictor of Gates MacGinitie Total Reading. PLS III had the highest Pearson correlation (r of .30) of any of the PLS or MAS subtests to the Gates MacGinitie Total Reading score. PLS III was noted by Werner (1992c) as an "auditory' subtest. The PLS III subtest tasks, upon more extensive informal examination, definitely appear to require an extensive cognitive component for successful completion.

Research Question Three

Research question three involved the extent to which the EPSF 7 developmental modality scores were related to and predicted future Gates MacGinitie reading achievement. All Pearson correlations were significant for the Gates MacGinitie Comprehension and Total reading scores to all the EPSF modality scores except the GM modality. All the EPSF modality scores except GM and EL were significantly Pearson correlated with the Gates MacGinitie Vocabulary scores. Thus GM was not significantly Pearson correlated with any of the three Gates MacGinitie reading scores. The most significant .01 level individual Pearson correlations for Gates MacGinitie Vocabulary, Comprehension and Total reading scores each consistently involved the RL modality with respective correlations of .22, .35 and .33. (The VD and FM modalites were also at the .33 correlation level with the Gates MacGinitie Total Reading score).

A correlation of .46 was found between the canonical data independent variable linear composite (i.e., the 7 EPSF modality scores) and the combined dependent variable of Gates MacGinitie Vocabulary and Comprehension scores. A shared variance of 21.38% was noted between the canonical combined 7 EPSF modality scores and the combined Gates MacGinitie Vocabulary and Comprehension variable set with all 7 EPSF modality scores except GM significant individual predictors (i.e., canonical loadings) of Gates MacGinitie combined Vocabulary and Comprehension scores.

The RL, AU and VM modality scores were currently noted to be the most significant individual canonical loadings with correlations of .73, .66 and .60, respectively. Bryant (1991) and Agostin (1993)also found (using the 7 EPSF developmentally modality scores) that RL, AU and VM in descending order were the most significant predictor of a future (locally developed) kindergarten test and Standard Achievement Test (SAT) reading achievement, respectively.

Research question three canonical redundancy coefficients found Gates MacGinitie combined Vocabulary and Comprehension scores accounted for or explained 60.42% of the variance in the 7 EPSF combined modality scores. Conversely, the EPSF 7 modality scores were able to account for or explain 32.92% of the variance in the Gates MacGinitie combined Vocabulary and Comprehension scores. These two same redundancy percentage scores are very similar to those generated when the 5 EPSF battery test scores were used as the defined combined canonical independent composite.

The stepwise multiple regression analysis for the 7 EPSF modality scores as predictors of the three individual Gates MacGinitie reading scores again found the pattern (as in research questions one and two) of lower overall total variance accounted for in Gates MacGinitie Vocabulary versus Gates MacGinitie Comprehension or Total Reading scores. The combined 7 EPSF modality scores accounted for only 6.6% of the total variance predicted in the Gates MacGinitie Vocabulary scores. The previous (research question two) 13 EPSF variable set (containing all EPSF tests and subtest) and the 11 variable EPSF set (compromised of all EPSF tests and subtests except the PLS Total and MAS Total) accounted for 8 % and 7.84% of the total variance In Gates MacGinitie Vocabulary scores. Thus, some noted increase in Gates MacGinitie

Vocabulary variance predicted (from 5.5% for the 5 EPSF tests alone) was noted when the EPSF modality or subtest scores were used as independent predictors.

The multiple regression analysis found the most significant EPSF modality individual predictors (or factor loadings) for Gates MacGinitie Vocabulary scores were RL and VM. Receptive language (RL) is noted in the literature as an excellent measure of vocabulary skills. Meta analysis research by Tramontana, et al. (1988) found that receptive language (RL), when used in a multimeasure assessment "was among the best predictors of first grade reading achievement", (p. 127). The EPSF modalities of RL and VM were noted to be the two highest Pearson correlated of the 7 EPSF modality scores to Gates MacGinitie Vocabulary scores. RL and VM were the best combined EPSF modality predictors (with 6.04% of total variance) noted in future Gates MacGinitie Vocabulary scores. In contrast, only 6.6% of the total variance in Gates MacGinitie Vocabulary scores could be accounted for or explained by all 7 EPSF modality scores. Thus, the deletion of the 5 EPSF modality scores of FM, AU, EL, GM and VD would not significantly decrease the amount of total variance predicted for Gates MacGinitie Vocabulary scores by the remaining RL and VM modalities.

The identical significant EPSF modality individual predictors of RL, FM, AU and VM were noted to account for 20.91% and 18.53% of the total variance in Gates MacGinitie Comprehension and Total Reading scores. Roth, et al. (1993) determined that the EPSF modality of FM was the best predictor of future grade retention, special education referral and eventual special education placement. Roth, et al. (1993) also noted the FM and AU modalities were significant predictors of future IOWA Tests of Basic Skills reading achievement. Tramontana, et al. ,1988 noted that visual motor (VM) proficiency was "among the best predictors of achievement in the first grade", (p. 128).

Overall, the addition of the other 3 EPSF modalites to the significant RL, FM, AU and VM would only raise the total variance predicted in Gates

MacGinitie Comprehension and Total Reading scores to 21.21% and 18.84%, respectively. Thus, the deletion of the EPSF modalities of EI, GM and VD would not significantly decrease the predictive capabilities for Gates MacGinitie Comprehension and Total Reading scores by the remaining 4 EPSF modalities. Tramontana, et al. (1988) noted the decreased importance of gross motor (GM) skills as a kindergarten predictor of future reading achievement.

Summary

Equal opportunity for all is a major tenet of the American way of life. Public schools try to reflect this philosophy in their educational programs. Early intervention and kindergarten screening programs are based on the assumption that early identification of potential learning problems can help identify children at risk of school failure and possible remove, through early intervention, potential roadblocks to the educational equality for all children.

No one can argue about the basic good intentions behind the development of kindergarten screening instrumentation. Still, despite the dramatic increase in kindergarten instrumentation generated in the last twenty five years, continued questions in the literature persist today regarding their psychometric properties and the practical utility of the wide array of kindergarten screening instrumentation now available. Some kindergarten screening instrument developers are quick to state that their instruments are merely a screening tool to identify children at risk. Thus, implying that their instrumentation is just one phase in the ongoing educational intervention process. But the fact remains American school systems today are increasingly introducing traditional academic skill training, especially in the areas of reading, writing and math, into the kindergarten environment (Charlesworth, 1989 and Slavin, et al. 1994). Thus, as the kindergarten environment increasingly becomes more accountable for academic skill acquisition, so should kindergarten screening developers be vigilant of the psychometric adequacies, purposes and applied uses of their instrumentation.

Meisels and other authorities have spoken for years about the increased confusion over the purposes and implementations of kindergarten screening instruments. Kindergarten screening instruments are still currently being misused for readiness as well as for diagnostic purposes. The EPSF begun in 1971 is one of the more notable kindergarten screening programs. It was initially developed with Title 1 funding with the stated intention of helping identify children at risk for school failure. It is interesting to note the EPSF program developers label their initial EPSF screening battery phase as the EPSF program "Diagnostic" component, (Werner, 1990). This EPSF terminology definitely does not help decrease the confusion that exists between kindergarten screening and diagnostic testing.

The EPSF screening battery and resultant 7 modality scores are generated for use in the EPSF program "curriculum design" and "classroom management " components that focus on reading and writing skill development. Meisels (1987) stated predictive validity is the major validity issue for screening tests due to their focus on learning potential. The EPSF screening battery must meet this defined predictive validity criteria for predicting future reading and writing success. This study examined the EPSF capabilities of the screening battery and derived modality scores to predict future first grade reading achievement as measured by the Gates MacGinitie Reading Tests (Third Edition). The Gates MacGinitie Reading Tests have been widely used as a reading screening instrument since 1926.

This study found that the basic 5 EPSF screening battery tests overall could only account for 5.5% of the predicted variance in Gates MacGinitie Vocabulary scores. It was found that the Gates MacGinitie Reading Test is more heavily weighted as a measure of Comprehension versus Vocabulary skills. Neither adding the 8 EPSF subtest scores to the 5 EPSF screening battery tests scores nor using the 7 EPSF modality scores as EPSF predictors could appreciably increase the amount of predicted Gates MacGinitie Vocabulary beyond 8%.

The canonical redundancy index found, given the Gates MacGinitie Vocabulary and Comprehension scores composite, that 32.92% of its variance could be accounted for or explained by the 5 EPSF screening tests canonical composite. Canonical analysis also noted the shared variance between the 5 combined EPSF screening tests and combined Vocabulary and Comprehension scores was 23.38%. The canonical loadings showed that the combined Gates MacGinitie Vocabulary and Comprehension composite had a correlation of .99 between the Comprehension subtest and the Vocabulary and Comprehension composite. Thus, the much higher shared variance for Gates MacGinitie combined Vocabulary and Comprehension scores (versus Vocabulary scores predicted alone) is more a reflection of the Comprehension subtest's influence.

The PPVT-R and VMI were the best combined EPSF screening basic tests battery predictors of Gates MacGinitie Vocabulary. The PPVT-R, VMI and PLS Total were Pearson correlated in the .30s to each other and all were the most significant combined predictors of Gates MacGinitie Comprehension and Total Reading scores. PPVT-R was the most significant individual predictor of Gates MacGinitie Vocabulary. VMI followed by PPVT-R and PLS Total were in identical order the best single (of the 5 basic EPSF screening tests) predictors of Gates MacGinitie Comprehension and Total reading scores.

Previous research has shown the potential for predicting reading achievement by the PPVT-R, VMI and PLS Total. McConnell (1986) and Terbush (1990) both noted that the EPSF tests of PPVT-R, PLS Total and VMI (in the same identical order) to be the best predictors the Metropolitan Readiness Tests and IOWA Tests of Basic Skills achievement, respectively. Fletcher and Satz (1982) in a seven year longitudinal follow-up for kindergarten prediction of reading achievement found the PPVT-R ,VMI and two non-standardized tests were the best kindergarten student predictors of academic risk seven years later at the end of sixth grade.

The MAS Total and DAP tests together only added to the total variance predicted by the combined PPVT-R, VMI and PLS Total less than 8% for Gates MacGinitie Vocabulary scores and 1% or less for the separate prediction of Gates MacGinitie Comprehension and Total Reading scores. The MAS Total and DAP were not significant combined or individual predictors of any of the three individual Gates MacGinitie reading scores. The MAS Total did (with a r of .31) just meet the minimal recommended .30 criterion as a significant canonical loading for prediction of the combined Gates MacGinitie Vocabulary and Comprehension composite. Still, the MAS Total and DAP could be dropped as potential predictors of Gates MacGinitie reading achievement with minimal loss of predictive capabilities by the remaining 3 EPSF basic screening battery tests of PPVT-R, VMI and PLS Total.

Present research found, when the EPSF 5 PLS subtests and 3 MAS subtests were examined, that none of the MAS subtests (as well as the MAS Total) were significant individual predictors of any Gates MacGinitie reading score. Werner (1992b) stated that the MAS scores "will be integrated with the outcomes of the total assessment, not interpreted as a single measure of performance in any one developmental area", (p. 16). Current research found the MAS Total and MAS 3 subtests were not be significant combined predictors (with other EPSF tests and/or subtests) of any of the three individual Gates MacGinitie reading achievement scores. In fact, MAS II was negatively correlated to Gates MacGinitie Comprehension and Total Reading performance with Pearson correlations of - .0393 and - .0176, respectively (see Appendix G). Also, MAS III had no significant correlation with any of the three Gates MacGinitie reading scores.

The MAS Total test and MAS subtests have been criticized previously in the literature by McConnell (1986) who found a "large number of students (in her study) receiving perfect MAS scores and almost no one scored at risk (for school failure)", (p. 91). McConnell (1986) stated the MAS Total test "appears to have inadequate ceiling to discriminate among children with varying motor skills", (p. 91) and "could be eliminated from the (EPSF) battery without appreciable loss of predictive power", (p. 91). Current research findings in this study support this contention. McConnell (1986) recommended further reliability and validity studies of the MAS due to, at that time, it was "poorly normed with no validity or reiiability data reported", (McConnell, 1986, p. 92).

The EPSF developers generated a 1992 EPSF.MAS Manual. It was commendable that the EPSF developers attempted to respond to the noted requests for published psychometric data on the MAS test and subtests. Still, it is interesting to note that the MAS Total test still relies on the original 28 items used since its inception. Thus, the MAS II subtest still has only 4 items and was currently found to be the least effective MAS subtest predictor of any of the three Gates MacGinitie reading scores.

Bracken (1987) commented that frequently preschool instrument generated subtest item gradients can be ineffective due to large change in children's obtained test scores caused by a single score. Test ceiling and limited test items are obvious psychometric concerns for the MAS II subtest. The EPSF test authors state the basic function of MAS II is to be used in conjunction with the VMI and DAP tests to form the FM developmental modality. Current data in this study found the MAS II to have nonsignificant Pearson correlations with FM, VMI and DAP of .00, .03 and .02, respectively.

The 1992 EPSF.MAS Manual is lacking in data presented to support reliability and validity claims made. The 1992 EPSF.MAS Manual report studies that state the MAS supposedly has adequate reliability as noted by a .90 interrater correlation (Crawford ,1989) and reported split half reliability coefficients of .58, .60 and .60 for MAS I, MAS II and MAS III (Thistlewaite & Cooke, 1992) as noted in Werner (1992b). But limited specific data was presented in the 1992 EPSF.MAS Manual to verify such claims. It is essential for all screening tests to meet the minimum test-retest and interrater reliability of at least .8 to .9. This information should be specifically reported in the MAS test manual and not just presented in general terms.

The 1992 reported MAS reliability and validity sample of 400 subjects did not meet recommended 1974 APA guidelines for 100 subjects per subset. The MAS 1992 reliability and validity sample had less than 100 subjects in 5 of the 6 defined age groups with only 26 and 29 subjects noted in the oldest and youngest age groups, respectively. This limited reliability and validity sample size is puzzling due to the vast number of potential EPSF subjects available to the EPSF test developers and their researchers. Compounding the problem is the fact this same "sample of 400 children from 10 school districts" (Werner, 1992b, p. 26) were reported in the 1992 EPSF.DAP Manual as used for its 1992 reliability and validity sample.

Other EPSF. MAS psychometric concerns include (1) the 1988 reported supposed "national " sample used to standardize the MAS test and (2) concern of the EPSF.MAS Manual reported validity claims. The 1992 EPSF.MAS Manual figuratively presented data showed that 33 of the 42 communities (or 78%) in the sample were in the states of Ohio and Illinois with 1684 of the total 3.093 children (or 54%) from the manual defined "North Central" region. Thus, more of a regional versus "National" 1988 standardization sample was used for the MAS test.

The 1992 EPSF.MAS Manual reported concurrent and construct validity information was minimal with no predictive validity information presented. Low construct validity correlations (based on child's age) ranging from .21 to .27 were found with no specific details or data to support construct validity claims made. Adequate MAS concurrent validity was professed in the 1992 EPSF.MAS Manual but not described adequately. Overall, the 1992 published psychometric properties of the MAS fall short of the mark in justifying its continued use.

The DAP was noted to not significantly contribute as a combined or individual predictor of any of the three individual Gates MacGinitie reading scores or the canonical combined Gates MacGinitie Vocabulary and Comprehension composite score. The DAP is used by the EPSF developers as a portion of the FM and VM developmental modalities when combined with component EPSF tests or subtests including VMI, MAS II and PLS V. The DAP was noted to have Pearson correlations for VMI, MAS II and PLS V ranging from .00 to .07 at best. In fact, the DAP had a Pearson correlation of .00 and .03 with the FM and VM modalities.

The 1992 EPSF.DAP Manual presents similar psychometric concerns as noted previously in this narrative regarding the EPSF.MAS. The 1990 DAP standardization sample reports no interrater reliability or judge scoring accuracy by the 79 different raters used in the national sampling. It is statistically imperative that a screening test such a the DAP which relies heavily on subjective scoring criteria must report at least a .8 interrater reliability coefficient as is recommended by the American Psychological Association (APA).

DAP internal consistency split-half coefficients (using the Kuder Richardson Formula 20) in the .70s were generally reported but not specifically described in the 1992 EPSF.DAP Manual. The 1992 MAS reliability and validity sample (of 400 children) psychometric limitation of not having 100 subjects per subset (in 5 of the 6 MAS subsets) is also relevant issue for the DAP which used the same set of subjects for its 1992 reliability and validity studies. The American Psychological Association (APA) even as early as 1974 stated that screening tests should have a minimum of 100 subjects for each screening test sample subgroup.

The noted concerns regarding the 1992 EPSF.DAP Manual reported construct validity correlations of only .22 to .27 for child's age involve the fact that specific detailed supportive data was not reported in the manual. "Relative modest correlations between the DAP and the PLS, MAS and PPVT " (Werner, 1992a, p. 17) were verbally reported as measures of DAP concurrent validity but were not disclosed in the 1992 EPSF.DAP Manual. The current study found Pearson correlations between DAP and PLS (Total), MAS(Total) and PPVT-R of .05, .01 and .10, respectively (see Table 1). Also, no statement of predictive validity data was presented in the 1992 EPSF.DAP Manual.

Overall, the DAP test does not meet the established APA psychometric criteria and consensus recommendations in the literature for acceptable reported reliability and validity limits for kindergarten screening tests. The current questionable psychometric qualities of the DAP along with the present study results, warrant the deletion of the DAP as a potential individual or combined predictor of Gates MacGinitie reading achievement.

The PLS Total subtests of PLS V and PLS II showed potential significant merit as possible combined predictors of Gates MacGinitie Vocabulary scores, if the PLS Total was not used as a EPSF predictor of Gates MacGinitie reading achievement. PLS III showed significant predictor capabilities as an individual predictor of Gates MacGinitie Total Reading score, if PLS Total was not used as a potential EPSF predictor of Gates MacGinitie reading achievement. PLS III is labeled as an "auditory" subtest but includes a large cognitive component as well as its EPSF defined "short term memory, association, sound discrimination and sequencing" (Werner, 1992c, p. 3) skills necessary for successful PLS III tasks completion. Previous EPSF developer research had found the PLS V, PLS III and PLS II (in descending order) to be the most noted PLS factors in the initial PLS test factor analysis. Still, very limited specific statistical data is presented in the PLS manual regarding the initial PLS test development factor analysis.

A critical point to be considered is that if the PLS factor analysis correlations are low, the PLS cannot predict any other test, much less the Gates MacGinitie. Pearson correlations from all 5 PLS subtests correlated at the .01 level of significance with PLS Total. All the PLS subtests were noted to be significant individual predictors of the combined Gates MacGinitie Vocabulary and Comprehension scores.

Overall, PLS Total remained a significant combined predictor of Gates MacGinitie Comprehension and Total Reading scores. PLS Total was noted to be the highest significant combined EPSF predictor (followed by PPVT-R and VMI) of Gates MacGinitie Comprehension scores (when all 13 EPSF tests and subtests or the 5 EPSF tests alone were considered as predictor variables). The best overall individual predictor of Gates MacGinitie Comprehension and Total Reading scores was VMI followed by PPVT-R and PLS Total. Current study data found continued inclusion of the PLS Total in the EPSF screening battery (along with the well established and psychometric sound PPVT-R and VMI) is justified as a potential significant combined predictor of Gates MacGinitie Comprehension and Total Reading scores.

The best overall combined and individual EPSF modality predictor of future Gates MacGinitie Vocabulary, Comprehension and Total Reading scores was RL. The EPSF modality of RL accounted for over one half of the noted significant total variance generated by the combined 7 EPSF modality scores. Receptive language (RL) has long been noted as a significant potential predictor of reading (especially in the area of vocabulary) by such noted researchers as Tramontana, et al. (1988) and Horn and Packard (1985). Previous EPSF modality research by Agostin (1993) and Bryant (1991) found RL to be the best predictor of a local generated kindergarten test and Stanford Achievement Tests (Reading scores), respectively.

Overall, the Gates MacGinitie Comprehension and Total Reading scores involved more individual and combined EPSF predictors than the Gates MacGinitie Vocabulary subtest. Pearson correlations showed the Gates MacGinitie Comprehension and Gates MacGinitie Total reading scores are correlated at .94. Thus, similar relationships were frequently noted between the Gates MacGinitie Comprehension and Total Reading scores in the current research.

Wallbrown, et al. (1975) noted that the Gates MacGinitie Comprehension subtest required a wider array of processing skills, especially in the visual motor area, than the Gates MacGinitie Vocabulary subtest. The increased significant of the combined predictors of RL, FM, AU and VM in the current Gates MacGinitie Comprehension and Total Reading scores could potentially attest to the increased processing skills required in reading comprehension versus vocabulary skills. The RL and VM were the only two combined current significant predictors of Gates MacGinitie Vocabulary scores.

The elimination of the EPSF modalities of GM, EL and VD would not significantly decrease the predictive capabilities of the remaining 4 EPSF modalities for Gates MacGinitie reading scores. Gross motor skills (GM) especially have been noted to be an ineffective kindergarten predictor of future reading performance (Tramontana, et al., 1988). The psychometric questionable DAP and MAS Total screening tests are two integral portions of the EPSF defined GM modality.

The best overall combined EPSF predictors of Gates MacGinitie Vocabulary scores (considering all the 5 EPSF tests , 8 EPSF subtests and 7 EPSF modality scores) were the RL and VM modalities with 6.04% of the predicted total variance in Gates MacGinitie Vocabulary. The PPVT-R and VMI followed closely behind RL and VM as the next best combined EPSF predictors of Vocabulary with a 5.08% of total variance predicted. The best single EPSF predictors of Gates MacGinitie Vocabulary (including all potential EPSF tests, subtests and modality scores) were RL and PPVT-R with 4.42% and 3.75% of total Gates MacGinitie Vocabulary variance predicted. PPVT-R is the major contributing EPSF test or subtest component to the formation of the RL modality. PPVT-R and RL have a Pearson correlation of .76 with a resultant significance level of .01 between them.

The best overall combined EPSF predictors of Gates MacGinitie Comprehension (considering all EPSF tests, subtest and modality scores) were PLS Total, VMI and PPVT-R accounting for a combined 23.18 % of the predicted variance in Gates MacGinitie Comprehension. PLS Total was the best combined predictor of Gates MacGinitie Comprehension accounting for 12% of the variance (followed by the RL modality with 11.3% of the variance) in Gates MacGinitie Comprehension scores.

The same three identical EPSF combined predictors of VMI, PPVT-R and PLS Total (noted in Gates MacGinitie Comprehension) were the best combined

EPSF predictors (considering all EPSF tests, subtests and modalities scores) of Gates MacGinitie Total Reading scores accounting for 22.35% of the variance in Gates MacGinitie Total Reading scores. VMI became the best combined EPSF predictor of Gates MacGinitie Total Reading scores accounting for 12.54% of the variance in Gates MacGinitie Total Reading scores.

Overall, the canonical analysis found that the combined 5 EPSF basic screening battery tests accounted for the most variance in the combined Gates MacGinitie Vocabulary and Comprehension subtests composite with 32.92% followed by the 7 EPSF modalities with 32.35% and the EPSF 5 tests/8 subtests (accounting for only 20.32%). Canonical correlations between the various EPSF variable sets were very closely related and ranged from .46 to .51. The highest shared variance between the various canonical EPSF composite variable sets and the Gates Vocabulary and Comprehension scores composite was the 5 EPSF tests and 8 subtests combined accounted for 25.85% (followed by the 5 EPSF tests alone with 23.38% and 7 EPSF modalites with 21.38%) of the shared variance between the two dependent and independent sets.

Conclusions

In conclusion, beyond the numerous ways to look at the current study results, there remains an underlying issue for the practitioner of which of the EPSF variables should be used to predict future overall Gates MacGinitie reading achievement. The basic EPSF screening battery tests of PPVT-R, VMI and PLS Total remain overall the best combined predictors of Gates MacGinitie reading achievement. All things considered, they represent the most efficient and effective EPSF screening variable to predict future Gates MacGinitie reading achievement. Deletion of the MAS Total and DAP tests from the basic EPSF screening battery would not significantly effect the predictive capabilities of the EPSF screening battery for Gates MacGinitie reading achievement . Also, the contribution of DAP and MAS Total as component parts of their respective

EPSF modalities scores is minimal.

The following consideratons are appropriate in light of previous and current research findings:

1. The EPSF basic screening battery tests of MAS and DAP need to be subjected to more rigorious psychometric exploration of their sample adequacies and tests item gradients (especially for MAS II). Further documentation and refinement of the MAS and DAP test developers' published reliability and validity studies is needed.

2. Continued EPSF staff research should include predictive validity studies of their basic screening battery.

3. It is recommended that the MAS and DAP screening battery tests not be considered as potential individual or combined (with other EPSF variables) predictors of future Gates MacGinitie reading achievement.

4. The EPSF staff stress the use of developmental modalities in their EPSF diagnostic screening battery and suggested supplimental classroom intervention lessons.. Further research is needed to clarify the effectiveness of developmental modality approaches in reading skill development.

5. Further kindergarten screening research on the three Gates MacGinitie reading scores as separate entities should be done to substantiate the potential different skills required for reading achievement.

6. Continued kindergarten screening research on prediction of the different development subskills required at different grade levels is needed.

7. Increased research stressing the practical utility of kindergarten screening instrumentation in identification of academically at-risk children is needed. Development of practical kindergarten screening instrument cutoff scores for predictive validity purposes is essential, if kindergarten screenings are to be effective in the identification of children at risk of future academic failure.

LITERATURE CITED

Adelman, H. S. (1982). Identifying learning problems at an early age: A critical appraisal. Journal of Clinical Child Psychology, 11 , 255-261.

Agostin, T. E. (1993). Predictive and discriminant validity of the social skills rating system and the early prevention of school failure program for first grader success. (Doctoral Dissertation, University of Mississippi, 1993), Dissertation Abstracts International, 54, 2883.

Anderson, K. C. (1985). Early prevention of school failure. Peotone, IL: Early Prevention of School Failure. (ERIC Document Reproduction Services No. ED 260 508)

Baenen, N, R, (1992). Early prevention of school failure: Longitudinal study 1987 - 1991: Evaluation report. Raleigh, NC: Wake County Public School System. (ERIC Document Reproduction Services No. ED 349 087)

Bailey, D. D. & Wolery, M. (1989). Assessing infants and preschoolers with handicaps. Columbus, OH: Charles E. Merrill.

Barnes, K. E. (1982). Preschool screening: The measurement and prediction of children at risk. Springfield, IL: Charles Thomas

Berry, K. E. (1989). Developmental Test of Visual - Motor Intergration manual (3rd ed.). Cleveland, OH: Modern Curriculum Press.

Betz, P. (1990). Evaluation of student impact and program fidelity: On the way to success in reading and writing. Early prevention of school failure. Peotone, IL: Early Prevention of School Failure.

Bloom, B. A. (1964). Stability and change in human characteristics. New York: John Wiley & Sons.

Bracken, B. A. (1987). Limitations of preschool instruments and standards for minimal levels of technical adequacy. Journal of Psychoeducational Assessment, 4, 313 - 326.

Brown, J. R. (1991). The retrograde motion of planets and children: Interpreting percentile rank. Psychology in the Schools, 28, 345 - 353.

Bruner, J. (1980). Under five in Britain. Ypsilanti, MI: High Scope Press.

Bryant, M. R. (1991). A discriminate analysis of early prevention of school failure preschool screening. (Doctoral Dissertation, University of Nevada at Reno, 1991). Dissertations Abstracts International, 52, 07A.

Butler,S. R., Marsh, H. 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.

Calfee, R. (1985). Gates - MacGinitie reading tests. Ninth Mental Mental Measurements Yearbook (pp. 593 - 595). Lincoln, NE: Buros Institute, University of Nebraska Press.

Catteral, J. & Cota-Robles, E. (1988). The educationally at- risk: What the numbers mean. Accelerating the education of at-risk student. Conference on Accelerating the Education of At-Risk Students (pp. 6 -7), Stanford, CA: Educational Research, Stanford University.

Charlesworth, R. (1989). "Behind" before they start? Deciding how to deal with the risk of kindergarten failure. Young Children, 44, 5 - 13.

Cook, N. R. & Smith, R. A. (1992). MAS reliability and validity: Analysis, synthesis and general discussion. In L. Werner (Ed.) Motor Activity Scale (pp. 28-36). Peotone, IL: Early Prevention of School Failure.

Cooter, R. B. Jr. & Curry, S. (1989). Gates MacGinitie Reading Tests, (3rd ed.). Reading Teacher, 43, 256 - 258.

Crawford, C. (1989). The effects of shadowing in perceptual-motor learning of kindergarten students. Unpublished doctoral dissertation, University of New Mexico, Albuquerque, NM.

deHirsch, K., Jansky, J. J. & Langford, W. S. (1966). Predicting reading failure. New York: Harper & Row.

Driscoll, W. T. (1992). Three year sustained effect of the early prevention of school failure on reading and language achievement. (Doctoral Dissertation, Columbia University Teachers College, 1992), Dissertation Abstracts International, 53, 1778.

Dunn, L. M. & Dunn, L. M. (1981). Peabody Picture Vocabulary Test - Revised manual. Circle Pines, MN: American Guidance Services.

Educational Programs that Work, (1994). Educational programs that work (EPTW): Catalogue of the national diffusion network (20th ed.). Longmont, CO: Sopris West.

Elkind, D. (1989). Miseducation - preschoolers at risk. New York: Alfred A. Knoff

Ellwein, M. C., Walsh, D. J. Eads, G, M. & Miller, A. K. (1991). Using readiness tests to route kindergarten students; The snarled intersection of psychometrics, policy and practice. Educational Evaluation and Policy Analysis, 13, 159 - 175.

Evans, R. & Ferguson, N. (1974). Screening school entrants. Association of Educational Psychologists Journal, 3, 2-9.

Fletcher, J. M. & Satz, P. (1982). Kindergarten prediction of reading achievement: A seven year longitudinal followup. Educational and Psychological Measurement, 42, 681 - 685.

Frechtling, J. A. (1989). Adminstrative uses of school testing programs. In R. Linn, (Ed.) Educational Measurement (3rd ed.), (pp. 475 - 487), NY: MacMillian.

Fredebaugh, D. (1984). The reliability of the Preschool Language Scale in St. Thomas. Unpublished master's thesis, College of the Virgin Islands, St. Thomas.

Garner, J. B. (1993). A multivariate model of the prediction of cognitive development, academic achievement and special instructional arrangement in school age children, (Doctoral Dissertation, Kent State University, 1993), Dissertation Abstracts International, 54, 3320.

Gary, J. E. (1975). An introduction to standardized testing for teachers and adminstrators. Framingham, MA: Educational Records Bureau. (ERIC Document Reproduction Services No. ED 117 1971)

Gracey, C. A., Azzara, C. V. & Reinherz, H. (1984). Screening revisited; A survey of U. S. requirements. Journal of Special Educaton, 18, 101 - 107.

Graham, L. M. (1990). Test review: Gates MacGinitie Reading Tests (3rd ed.). Reading Improvement, 27 , 21 - 23.

Graue, M. E. (1993). Ready for what? Constructing meanings of readiness for kindergarten. Albany, NY: State University of New York Press.

Green, D. R. (1987). A guide for interpreting standardized test scores. NASSP Bulletin, 71 , 23 - 25.

Greenfield, D. B. & Scott, M. S. (1985). A cognitive approach to preschool screening. Learning Disabilities Research, 1 , 42 - 49.

Gridley, B. E., Mucha, L. & Hatfield, B. B. (1995). Best practices in preschool screening. In A. Thomas & J. Grimes, (Eds.) Best practices in school psychology (pp. 213 - 225). Washington, DC: National Association of School Psychologists.

Hanna, G. S., Dyck, N. S. & Holen, M. C, (1980). Comparing potential with achievement: Rationale and procedures for objectively analyzing achievement - aptitude discrepancies in LD classification. Kansas State University; Manhattan, Kansas. (ERIC Document Reproduction Services No. ED 195 095)

Harrington, R. G. (1984). Preschool screening: The school psychologist's perspective. School Psychology Review, 13, 363 - 374.

Harrison, P. L. (1993 October). Roles of school psychologists in preschool screening and assessment programs. Paper presented at the meeting of the Oklahoma School Psychological Association, Tulsa, OK.

Horn, W. F. & Packard, T. (1985). Early identification of learning problems; A meta - analysis. Journal of Educational Psychology, 77, 597 - 607.

Joiner, L. M. (1977). A technical analysis of the variation in screening instruments and programs in New York State. City University of New York, NY Center for Advanced Study in Education. (ERIC Document Reproduction Services No. ED 154 596)

Kelley, M. F. & Surbeck, E. (1983). History of preschool assessment. In K. D. Paget & B. A. Bracken, (Eds.). The psychoeducational assessment of preschool children (pp. 1-16). NY: Grune & Stratton.

Keogh, B. K. & Becker, L. (1973). Early detection of learning problems: Questions, cautions and guidelines. Exceptional Children, 40, 5-11.

Lehr, C. A., Ysseldyke, J. E. & Thurlow, M. L. (1987). Assessment practices in model early-childhood special education programs. Psychology in the Schools, 24, 390 - 399.

Levin, H.M. (1985). The educationally disadvantaged: A national crisis. The State Youth Initaitives Project, Working Paper #6. Philadelphia: Public/Private Ventures.

Lichenstein, R. & Ireton, H. (1984). Preschool screening: Identifying young children with developmental and educational problems. Orlando, FL:

Grune & Stratton.

Lindquist, G. T. (1982). Preschool screening as a means of predicting later reading achievement. Journal of Learning Disabilities, 15, 331 - 338.

Lindsay, G. A. & Wedell, K. (1982). The early identification of educationally "at - risk" children revisited. Journal of Learning Disabilities, 15, 212 - 217.

MacGinitie, W. H. & MacGinitie, R. K. (1989a). Gates-MacGinitie Reading tests, third edition. Reading Teacher, 43, 256 - 258.

MacGinitie, W. H. & MacGinitie, R. K. (1989b). Gates- MacGinitie Reading Tests, third edition: Manual for scoring and interpretation. Chicago, IL: Riverside.

MacGinitie, W. H. & MacGinitie, R. K. (1989c). Gates-MacGinitie Reading Tests (third edition): Technical report. Chicago, Il: Riverside.

McConnell, S. S. (1986). Prediction and prevention of learning difficulties among kindergarten students. (Doctoral Dissertation, Ball State University, 1986), Dissertation Abstracts International, 47, 2089.

McGowen, S. (1991). Relationship among scores on the Bracken Basic Concept Scale and the Differential Ability Scales with a preschool sample. Unpublished master's thesis, Oklahoma State University, Stillwater, OK.

Mcloughlin, C. S. & Rausch, E. (1990). Best practices in kindergarten screening. In A. Thomas & J. Grimes (Eds.) Best practices in school psychology II (pp. 455 - 467). Washington, DC: National Association of School Psychologist.

Meisels, S. J. (1985). Developmental screening in early childhood: A guide. Washington, DC: National Association for the Education of Young Children.

Meisels, S. J. (1986). Testing four and five - year -olds: Response to Salzer and to Shephard and Smith. Educational Leadership, 44, 90 - 92.

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

Meisels, S. J. (1989). Can developmental screening tests identify children who are developmentally at risk? Pediatrics, 83, 578 - 585.

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

Mercer, C. D., Algozzine, B. & Trifiletti, J. (1979). Early identification: An analysis of the research. Learning Disabilities Quarterly, 2, 12 - 24.

Miller, L.J. & Schouten, G. W. (1988). Age-related effects on the predictive validity of the Miller Assessment for Preschoolers. Journal of Psychoeducational Assessement, 6, 99-106.

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

Nuttal, E., Romero, I. & Kalesnik, J. (1992). Assessing and screening preschoolers: Psychological and educational dimensions. Needham Heights, MA: Allyn & Bacon.

Olson, L. (1991). Social woes pose threat to reform: C.E.D. maintains. Education Week, 24, 1.

Paget, K. D. (1990). Best practices in assessment of competence in preschool - age children. In A. Thomas & J. Grimes (Eds.) Best Practices in School Psychology II (pp. 107 - 119). Washington, DC: National Association of School Psychologists.

Paget, K. D. & Nagle, R. J. (1986). A conceptual model of preschool assessment. School Psychology Review, 15, 154 - 165.

Parker, L. D. & Clechalski, J. C. (1990). Kindergarten screening and remedial strategies with implications for school counselors. Elementary School Guidance and Counseling, 25, 117 - 122.

Patrick, A. H., Kimball, G. H. & Crawford, J. (1984). Qualitative and quantitive evaluation of the Early Prevention of School Failure Program for kindergarten students. Planning, Research and Evaluation Department, Oklahoma City Public Schools, Oklahoma City, OK. (ERIC Document Reproduction Services No. ED 244 981)

Pedhuzur, E. J. (1982). Multiple regression in behavioral research: Explanation and prediction (2nd ed.). New York: Holt, Rinehart & Winston.

Phillips, S. E. & Clarizio, H. (1988). Limitations of standard scores in individual achievement testing. Educational Measurement: Issues and Practice, 7 , 8 - 15.

Piaget, J. (1952). The origins of intelligence in children. (M. Cook, Trans.) New York: International University Press.

Repeating grades in school: Current practice and research evidence (1990). (Report No. EA 021686). New Jersey: Rutgers, Center for Policy Research in Education. (Eric Document Reproduction Services No. ED 323 585)

Reynolds, C. D. & Kamphaus, R. W. (1990). Handbook of psychological and educational assessment of children: Intelligence and achievement. New York: Guilford.

Rosenkoetter, S. & Wanska, S. (1992). Best Practices in Preschool Screening. Paper presented at the 70th Annual Convention of the Council for Exceptional Children, Baltimore, MD. (ERIC Document Reproduction Services No. ED 347 748)

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

Rudner, L. M. (1989). Basic testing principles. In L. M. Rudner, J. C. Conoley, & B.S.Plake, (Eds.) Understanding achievement tests: A guide for school adminstrators. (pp. 9 - 52). Washington, DC: ERIC Clearinghouse on Tests, Measurement and Evaluation.

Salvia, J. & Ysseldyke, J. E. (1991). Assessment (5th ed.). Boston, MA: Houghton - Mifflin.

Sattler, J. M. (1982). Assessment of children's intelligence and special abilities. Boston: Allyn & Bacon.

Sattler, J. M. (1990). Assessment of children (3rd ed.). San Diego, CA: Jerome M. Sattler.

Satz, P. & Fletcher. J. (1979). Early screening tests: Some uses and abuses. Journal of Learning Disabilities, 12, 43 -50.

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

Shephard, L. A. & Smith, M. L. (1986). Synthesis of research on school readiness and kindergarten retention. Educational Leadership, 44, 78 -

85.
Shephard, L. A. & Smith, M. L. (1988). Escalating academic demand in kindergarten: Counterproductive polices. Elementary School Journal, 89, 135 - 144.

Slavin, R. E., Karweit, N. L. & Madden, N. A. (1989). Effective programs for students at risk. Needham Heights, MA: Allyn & Bacon.

Slavin, R. E., Karweit, N. L. & Wasik, B. A. (1994). Preventing school failure: research, policy and practice. Boston, MA: Allyn & Bacon.

Slavin, R. E. & Madden, N. A. (1989). What works for students at risk: A research synthesis. Educational Leadership, 46, 4 - 13.

Stangler, S. R., Huber, C. J. & Routh, D. K. (1980). Screening growth and development of preschool children: A guide for test selection. New York: McGraw - Hill.

Steven, J. (1992). Applied multivariate statistics for the social sciences (2nd ed.). Hilldale, NJ: Lawrence Erlbaum Associates.

Strand, K. & Werner, L. (1981). Evaluation summary of the Early Prevention of School Failure Project: A national validation program 1971 - 1981. Peotone, IL: Early Prevention of School Failure Project. (ERIC Document Reproduction Services No. ED 331 630)

Talmadge, G. K. (1976). Interpreting NCEs. ESEA Title I Evaluation and Reporting System. Mountain View, CA: RMC Research Corporation.

Terbush, R. L. (1990). Analysis of selected premeasures and achievement of kindergarten and first - grade students. (Doctoral Dissertation, Northern Arizona University, 1990). Dissertation Abstracts International, 51, 1589.

Terbush, R. L., Bliss, S. W. Staines, E. H., Deneshinsky, M. I. & Dankard, R. D. (1990). The relationship between selected skill measurements of kindergarten and first grade students and academic success. Paper presented to National Council of States on Inservice Education. Orlando, FL. (ERIC Document Reproduction Services No. ED 331 630)

Thompson, B. (1984). Canonical correlation analysis: uses and interpretation. Beverly Hills, CA: Sage Publications.

Thurlow, M. A., O'Sullivan, P. J. & Ysseldyke, J. E. (1986). Early screening for special education: How accurate? Educational Leadership, 44, 93 - 95.

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

Wallbrown, J. D., Engin, A. W., Wallbrown, F. H. & Blaha, J. (1975). The prediction of first grade reading achievement with selected perceptual - cognitive tests. Psychology in the Schools, 12, 140 - 149.

Ward, J. G. & Gould, J. C. (1980). Plain talk about standardized tests. Washington, DC: American Federation of Teachers. (ERIC Document Reproduction Services No. ED 206 721)

Werner, L. (1990). Early Prevention of School Failure training and resource book: Screening & conferencing for educational planning (4th ed.). Peotone, IL: Early Prevention of School Failure.

Werner, L. (1991). Early Prevention of School Failure: Awareness packet. Peotone, IL: Early Prevention of School Failure.

Werner, L. (1992a). The Early Prevention of School Failure Draw A Person notebook. Peotone, IL: Early Prevention of School Failure.

Werner, L. (1992b). Motor Activity Scale. Peotone, IL: Early Prevention of School Failure.

Werner, L. (1992c). Preschool Language Scale. Peotone, IL: Early Prevention of School Failure.

Ysseldyke, J. E. Thurlow, M. A., O'Sullivan, P. J. & Bursaw, R. A. (1986). Current screening and diagnostic practices in a state offering free preschool screening since 1977: Implications for the field. Journal of Psychoeducational Assessment, 4, 191 - 201.

Zeh, J. D. & Baenen, N. R. (1991). Early Prevention of School Failure evaluation report. Wake County Public School System, Raleigh, NC. (ERIC Document Reproduction Service No. ED 351 097)

Zeitlin, S. (1976). Kindergarten screening: Early identification of potential high-risk learners. Springfield, IL: Charles C. Thomas.

APPENDIXES

APPENDIX A

EPSF GENERATED SCREENING PROFILE

(Sample Teacher Pretest Report)

EPSF DIAGNOSTIC STUDENT PROFILE
Chatham

STUDENT NAME: Sarah Moyer
DOB: 8/15/84
CHRONOLOGICAL AGE: 5-1

TEACHERS NAME: Mrs. Lewis
TEST DATE: 9/1/89

	RL	EL	AD	VM	VD	FM	GM
CONSIDERABLE STRENGTH	.	.	X	.	.	.	.
MODERATE STRENGTH	.	.	.	.	.	.	.
AVERAGE	.	X	.	.	.	.	X
MODERATE NEED	X	.	.	X	X	X	.
CONSIDERABLE NEED	.	.	.	.	.	.	.

RL = Receptive Language
EL = Expressive Language
AD = Auditory
VM = Visual Memory
VD = Visual Discrimination
FM = Fine Motor
GM = Gross Motor

SPEECH OBSERVATIONS:
ARTICULATION: CLEAR
HEARING: Passed
VISION: Passed

RECEPTIVE LANGUAGE: Moderate Need

EXPRESSIVE LANGUAGE: Average

AUDITORY: Considerable Strength

VISUAL MEMORY: Moderate Need

VISUAL DISCRIMINATION: Moderate Need

FINE MOTOR: Moderate Need

GROSS MOTOR: Average

SUPPORT INFORMATION:
Sarah was able to identify the following colors:
red blue green orange yellow white
Sarah was able to identify the following shapes: triangle circle
Sarah could count to 9 in sequence.
Sarah WAS able to print her name.
Lateral dominance was as follows: FOOT = R HAND = L EYE = R

RESULTS:
PPVT = 4 - 3 VMI = 3 - 11 DAP = 4 - 0 MAS = AV
PLS I = MN PLS I = MS PLS III = CS PLS IV = AV PLS V = MN

COMMENTS:
PLS comment: Followed two auditory and two visual directions/Good listening skills
PPVT comment: Slow to respond/Repeated word
VMI comment: Task difficult/Switched hands
DAP comment: Poor pencil grip/Few details
MAS comment: Good balance and dexterity

APPENDIX B

1993 EPSF PROJECT DIRECTOR WRITTEN CORRESPONDENCE REGARDING 1992-93 LONGITUDINAL STUDY OF EPSF TREATMENT EFFECTS RESEARCH (using the Gates MacGinitie Reading Tests)

On the Way to SUCCESS in Reading and Writing
with
EARLY PREVENTION OF SCHOOL FAILURE
Nationally Validated Program

Luceille Werner
National Director

March 2, 1993

Mr. Greg Reed
2924 Crescent Drive
Stillwater, OK 74075

Dear Mr. Reed:

I spoke with your wife in Corpus Christi and was pleased to learn you are working on your doctoral dissertation related to the Early Prevention of School Failure Program. We want to assist you with any information you may need for your study. It is very important that following the assessment of the children that the classroom teacher and other resource staff provide daily treatment in area of need.

I am enclosing a copy of our new national longitudinal study. Perhaps you would like to be involved in assisting us with part of our study as a comparison.

Under separate cover I am sending you copies of the three assessment manuals. Would you please share the manuals with the Stillwater staff? Look forward to keeping in touch.

Phyllis Betz, 296 Laurel Park Place, Hendersonville, North Carolina, (704)692-9895, coordinates all our program studies and works with the outside university evaluator. Feel free to contact her if you have questions.

Sincerely,

Luceille Werner

Luceille Werner
National Director

LW/dd

Enclosure

cc: Phyllis Betz, Project Evaluator

Debra Murphy, Oklahoma State Facilitator

114 North Second Street • P.O. Box 956 • Peotone, Illinois 60468
708/258-3478 • 800/933-3478 • FAX 708/258-3484

On the Way to Success in Reading and Writing with Early Prevention of School Failure
Summary of Longitudinal Research Methodology (1992-96)

EPSF Kindergarten with Follow-up in SUCCESS First Grade and Second Grade (Sustained Effects)

1. **Design.**
   a. Timing. Instruments will be administered in 1992, kindergarten; 1993, post-kindergarten; 1994, post-first grade; and 1995, post-second grade to all program groups and the comparison groups.

   b. Groups. Kindergarten and first grade at-risk students will receive daily supplementary small group instruction in areas of identified need. The comparison group will receive the regular primary grade program at school site.

   c. Standard of Comparison. The standard of comparison will be the qualitative and quantitative differences between the comparison group of students, all of whom received regular classroom instructional programs and the program group. In addition, the program groups and the comparison groups will be compared with the national norm groups for the Gates-MacGinitie (quantitative analysis) and the SUCCESS Reading and Writing Checklists (qualitative anaysis).

2. **Sample.** Within each district, program and comparison groups will be selected based upon similarity of student populations on relevant educational characteristics. Students were selected for program and comparison groups in exactly the same manner that adopters use to identify students at risk, i.e., two or more years developmentally delayed in three or more of seven developmental areas (visual discrimination, visual memory, expressive language, receptive language, auditory, fine and gross motor). Comparison group teachers will not have access to staff development or curriculum materials and will not participate in workshops. Because the program is designed for use with <u>all</u> kindergarten and first grade populations, districts will be selected for their diversity <u>not</u> similarity of demographic characteristics. Students with considerable and moderate developmental needs will be targeted for the study because they are most affected by inappropriate academic programs regardless of other variables. Socio-economically diverse school districts will participate in the evaluation. The comparison and program samples, comprised of students who are identified as one or more years developmentally delayed, will provide a good representation of the program's intended at-risk target population and of program participants in replicating districts throughout the nation.

3. **Instruments and Procedures.** The Gates-MacGinitie instrument was chosen because it was judged by program developers and independent evaluators to be a valid standard measure of intended program outcomes; has been standardized relatively recently; and has derived standard scores for cross-comparison. In addition, reported results of reliability and validity studies are good. The SUCCESS (1) Reading and (2) Writing checklists were developed to provide an authentic measure of student progress. **The ecological approach to the study will recognize that the educational process is affected by the organization of the school and the classroom; the environmental history of children and their families; and the community aspects that affect schooling norms.**

4. **Data Collection.** The instruments will be administered prior to program instruction in the fall of 1992 (pretest) and again following completion of the kindergarten program in 1993; the first grade program in 1994; and the second grade regular classroom instructional program in 1995 to the sample of program students and the comparison groups.

   Every attempt will be made to conform exactly to the Gates-MacGinitie Test administration rules. Instructions, practice, problems, timing, scoring procedures, examiner qualifications, student assembly, and related concerns will be identified and communicated to certified program trainers, local program administrators, and/or experienced professional staff who will be trained to function as **in-school evaluators** to insure collection of accurate and complete data. To facilitate scoring and interpretation, a form has been designed to incorporate test dates, group identification numbers of anonymity, uniform entries for appropriate test levels, school and student demographics. Comparison groups and program groups will be identified as high or moderate risk students in kindergarten using the EPSF assessment battery (PPVT, VMI, DAP, PLS, MAS) in 1992. Program student groups and comparison groups will be identified as those functioning one or more years developmentally below their chronological ages in three or more of seven developmental areas.

5. **Analysis** of covariance (ANCOVA) will be used to compare treatment and comparison groups. Qualitative outcomes will be analyzed and compared with quantitative data.

10/19/92

APPENDIX C

PEARSON INTERCORRELATIONAL MATRIX OF THE 5 EPSF SCREENING TESTS

APPENDIX C

Pearson Intercorrelation Matrix of the 5 EPSF Screening Tests

Variable	PPVT-R	VMI	PLS Total	MAS Total
VMI	.0897			
PLS Total	.5770 **	.1693 **		
MAS Total	.0147	.0394	.4644 **	
DAP	.0993	.0750	.0453	.1015

N = 373

** = $p < .01$

APPENDIX D

PEARSON INTERCORRELATIONAL MATRIX OF THE 5 EPSF SCREENING TESTS AND 3 GATES MACGINITIE READING SCORES

APPENDIX D

Pearson Intercorrelational Matrix of the 5 EPSF Screening Battery Tests
and 3 Gates MacGinitie Reading Scores

EPSF Variable	Gates MacGinitie Reading Scores in Vocabulary	Comprehension	Total Reading
PPVT-R	.1937 **	.3446 **	.3150 **
VMI	. 1322 *	. 3331 **	.3542 **
PLS Total	. 1559 **	. 3516 **	.3263 **
MAS Total	. 1207 *	.1499 **	. 1619 **
DAP	. 0108	. 0914	. 0808

N= 373

** = $p < .01$

* = $p < .05$

APPENDIX E

CANONICAL FUNCTION COEFFICIENTS OF THE 5 EPSF TEST SCORES AND GATES MACGINITIE VOCABULARY AND COMPREHENSION SCORES

APPENDIX E

Canonical Function Coefficients of the 5 EPSF Test Scores
and Gates MacGinitie Vocabulary and Comprehension Scores

Independent Variable		Dependent Variable	
PPVT-R	.44	Gates MacGinitie Vocabulary	.02
VMI	.58	Gates MacGinitie Comprehension	.99
PLS Total	.39		
MAS Total	.04		
DAP	.09		

N = 373

APPENDIX F

PEARSON CORRELATIONS BETWEEN THE GATES MACGINITIE TOTAL READING, VOCABULARY AND COMPREHENSION SCORES

APPENDIX F

Pearson correlations between the Gates MacGinitie
Total Reading, Vocabulary and Comprehension Scores

Variable	Vocabulary	Comprehension	Total Score
Vocabulary	1.0		
Comprehension	.4332 **	1.0	
Total Score	.4726 **	.9319 **	1.0

N = 373
** = $p < .05$

Variable	Vocabulary	Comprehension	Total Score
Vocabulary	1.0		
Comprehension	.4769 **	1.0	
Total Score	.5283 **	.9346**	1.0

N = 630
** = $p < .01$

APPENDIX G

PEARSON CORRELATIONAL MATRIX
OF THE 8 EPSF SCREENING BATTERY SUBTESTS
AND 3 GATES MACGINITIE READING SCORES

APPENDIX G

Pearson Intercorrelational Matrix of the 8 EPSF Screening Battery Subtests and 3 Gates MacGinitie Reading Scores

EPSF	Gates MacGinitie Reading Scores in:		
Variable	Vocabulary	Comprehension	Total Score
PLS I	.1074 *	.2299 **	.2232 **
PLS II	.0259 *	.1753 **	.1725 **
PLS III	.1350 **	.3093 **	.2985 **
PLS IV	.0437	.2038 **	.1840 **
PLS V	.1512 **	.2218 **	.1911 **
MAS I	.1039 *	.1270 *	.1331 *
MAS II	.0101	-.0393	-.0176
MAS III	.0784	.1048	.1129

N= 373

** = $p < .01$

* = $p < .05$

APPENDIX H

CANONICAL FUNCTION COEFFICIENTS OF THE 5 EPSF SCREENING BATTERY TESTS AND 8 EPSF SUBTEST SCORES AND THE GATES MACGINITIE VOCABULARY AND COMPREHENSION SUBTEST SCORES

APPENDIX H

Canonical Function Coefficients of the 5 EPSF Screening Battery Tests and 8 EPSF Subtests Scores and the Gates MacGinitie Vocabulary and Comprehension Subtest Scores

Independent Variables	
PPVT-R	.44
VMI	.58
PLS Total	8.71
PLS I	2.32
PLS II	2.06
PLS III	2.91
PLS IV	2.47
PLS V	2.23
MAS Total	.10
MAS I	-.18
MAS II	-.10
MAS III	-.01
DAP	.08

Dependent Variables	
Gates MacGinitie Vocabulary	.04
Gates MacGinitie Comprehension	.98

APPENDIX I

SUPPLEMENTAL MULTIPLE REGRESSION ANALYSIS
TO DETERMINE THE TOTAL AMOUNT OF VARIANCE IN
GATES MACGINITIE READING ACHIEVEMENT AS
PREDICTED BY 13, 11, OR 8 EPSF
INDEPENDENT VARIABLE SETS

APPENDIX I

Supplemental Multiple Regression Analysis to determine the
Total Amount of Variance in Gates MacGinitie Reading Achievement
as Predicted by 13,11, or 8 EPSF Independent Variable Sets

No. of Variables	EPSF Variable Sets	Gates MacGinitie Reading Scores in Vocabulary	Comprehension	Total Score
5	5 Tests *	.0550	.2338	.2246
13	5 Tests & 8 Subtests **	.0800	.2313	.2443
11	3 tests & 8 Subtests ***	.0784	.2418	.2355
8	8 subtests only ****	.0442	.1467	.1280

N = 373

* = PPVT-R, VMI, PLS Total, MAS Total & DAP

** = PPVT-R, VMI, PLS Total, MAS Total, DAP, PLS I, PLS II, PLS III, PLS IV, PLS V, MAS I, MAS II, MAS III

*** = PPVT-R, VMI, DAP, PLS I, PLS II, PLS III, PLS IV, PLS V, MAS I, MAS II, MAS III

**** = PLS I, PLS II, PLS III, PLS IV, PLS V, MAS I, MAS II, MAS III

APPENDIX J

SUPPLEMENTAL MULTIPLE REGRESSION COMPARISONS OF THE MOST SIGNIFICANT 13, 11, OR 8 EPSF INDEPENDENT VARIABLE SET PREDICTORS (I.E. SIGN. IN F CHANGE) OF GATES MACGINITIE READING ACHIEVEMENT

APPENDIX J

Supplemental Multiple Regression Comparisons of the most Significant 13, 11, or 8 EPSF Independent Variable Set Predictors (i.e. sign. in F change) of Gates MacGinite Reading Achievement

EPSF Variable Set	Gates MacGinitie Reading Test Scores: Vocabulary			Comprehension			Total Reading		
	Variable	R^2 Change	Sign. F Ch.	Variable	R^2 Change	Sign F Ch.	Variable	R^2 Change	Sign F Ch.
13 Variables *	PPVTR	.0349	.000	PLS Total	.1200	.000	VMI	.1254	.000
	PLS 5	.0469	.018	VMI	.1960	.000	PPVTR	.2063	.000
	PLS 2	.0543	.050	PPVTR	.2250	.000	PLS 3	.2243	.000
11 Variables **	PPVTR	.0349	.000	PPVTR	.1164	.000	VMI	.1231	.000
	PLS 5	.0469	.018	VMI	.2066	.000	PPVTR	.2020	.000
	PLS 2	.0543	.050	PLS 3	.2235	.003	PLS 3	.2180	.004
				PLS 5	.2295	.049			
8 Variables ***	PLS 5	.0262	.003	PLS 3	.0957	.000	PLS 3	.0866	.000
	PLS 3	.0291	.037	PLS 5	.1224	.000	PLS 1	.1038	.005
				PLS 1	.1296	.014			

N = 373

* = PPVT-R, VMI, PLS Total, DAP, MAS Total, PLS I, PLS II, PLS III, PLS IV, PLS V, MAS I, MAS II, MAS III

** = PPVT-R, VMI, DAP, PLS I, PLS II, PLS III, PLS IV, PLS V, MAS I, MAS II, MAS III

*** = PLS I, PLS II, PLS III, PLS IV, PLS V, MAS I, MAS II, MAS III

APPENDIX K

CANONICAL FUNCTION COEFFICIENTS OF THE 7 EPSF MODALITY SCORES AND GATES MACGINITIE VOCABULARY AND COMPREHENSION SCORES

APPENDIX K

Canonical Function Coefficients of the 7 EPSF Modality Scores and Gates MacGinitie Vocabulary and Comprehension Scores

Independent Variable		Dependent Variable	
RL	.47	Gates MacGinitie Vocabulary	.06
EL	.10	Gates MacGinitie Comprehension	.97
AU	.36		
VM	.27		
VD	.13		
FM	.22		
GM	.04		

APPENDIX L

INSTITUTIONAL REVIEW BOARD APPROVAL

OKLAHOMA STATE UNIVERSITY
INSTITUTIONAL REVIEW BOARD
HUMAN SUBJECTS REVIEW

Date: 03-31-95 **IRB#:** ED-95-061

Proposal Title: PREDICTIVE ANALYSIS OF EPSF KINDERGARTEN MEASURES TO GATES MACGINITIE READING TESTS PERFORMANCE IN FIRST GRADE STUDENTS

Principal Investigator(s): Kay Bull, Gregory Wayne Reed

Reviewed and Processed as: Exempt

Approval Status Recommended by Reviewer(s): Approved

ALL APPROVALS MAY BE SUBJECT TO REVIEW BY FULL INSTITUTIONAL REVIEW BOARD AT NEXT MEETING.
APPROVAL STATUS PERIOD VALID FOR ONE CALENDAR YEAR AFTER WHICH A CONTINUATION OR RENEWAL REQUEST IS REQUIRED TO BE SUBMITTED FOR BOARD APPROVAL.
ANY MODIFICATIONS TO APPROVED PROJECT MUST ALSO BE SUBMITTED FOR APPROVAL.

Comments, Modifications/Conditions for Approval or Reasons for Deferral or Disapproval are as follows:

Signature: John H. Wyckoff Date: April 5, 1995
Chair of Institutional Review Board

VITA

Gregory Wayne Reed

Candidate for the Degree of

Doctor of Philosophy

Thesis: PREDICTIVE ANALYSIS OF EPSF KINDERGARTEN SCREENING MEASURES TO GATES MACGINITIE READING TESTS PERFORMANCE IN FIRST GRADE STUDENTS

Major Field: Applied Behavioral Studies

Biographical:

Personal Data: Born in Detroit, Michigan. Currently residing in Stillwater, Oklahoma. Married to Dr. Deborah Reed. Children: Christopher Price, Stephanie Kay and Amanda Michelle.

Education: Watonga High School, Watonga, Oklahoma graduated in May, 1964. Bachelor of Science degree in Psychology from Southwestern Oklahoma State University, Weatherford, Oklahoma in December, 1973. Masters of Education degree from Southwestern Oklahoma State University, Weatherford, Oklahoma in July, 1974. Completed the requirements for the Doctor of Philosophy at Oklahoma State University, Stillwater, Oklahoma in December, 1995.

Professional: Oklahoma licensed Marital and Family Therapist #344. Oklahoma licensed Licensed Professional Counselor #508. Oklahoma certifications as School Psychologist and School Psychometrist. Charter member of the Oklahoma School Psychological Association. Twenty years experience in private and public agencies as a staff psychologist. Currently employed by the Oklahoma State Department of Health (Child and Behavioral Health Division) as Psychologist and Director of Child Guidance Services in Noble and Pawnee Counties.