12-2005

Evaluating the Predictive Validity of the Speed DIAL Version of the DIAL-3, Developmental Indicators for the Assessment of Learning.

Roger Anthony Walk
East Tennessee State University

Evaluating the Predictive Validity of the Speed DIAL Version Of the DIAL-3, Developmental Indicators for the Assessment of Learning

---

A dissertation

presented to

the faculty of the Department of Educational Leadership and Policy Analysis

East Tennessee State University

In partial fulfillment

of the requirements for the degree

Doctor of Education

---

by

Roger Anthony Walk

December 2005

Dr. Louise MacKay, Chair

Dr. Hal Knight

Dr. Elizabeth Ralston

Dr. Terrence Tollefson

Keywords: Speed DIAL, DIAL-3, preschool screening, predictive validity, correlation, Virginia Standards of Learning Tests

ABSTRACT

Evaluating the Predictive Validity of the Speed DIAL Version Of the DIAL-3, Developmental Indicators for the Assessment of Learning

by

Roger Anthony Walk

The primary purpose of this research was to validate psychometric integrity of the Speed DIAL screening instrument by establishing its predictive validity. The validation process included determining what, if any, associations existed between the predictor variable, Speed DIAL and the outcome variable, the Virginia Standards of Learning third-grade math and total English tests.

Gender and age were factors in the study. Finally, certain intervening variables, those occurring after the predictor variable, but before the outcome variable, were included in the study. These interventions were supplementary programs or placements provided students in the hopes of positively influencing students' academic performance. The documented interventions were: retention status, special education status, number of years participating in Title I math, number of years participating in Title I reading, and number of years participating in the Phonological Awareness Literacy Screening (PALS) program.

Based on the findings, Speed DIAL does possess predictive qualities. There was a moderate correlation between Speed DIAL scores and Virginia Standards of Learning third-grade math and total English test scores. Speed DIAL's overall effectiveness rating exceeded 75%. Females scored higher than males on the Speed DIAL, and there was a negative association between Speed DIAL and the documented intervening variables. Using the Elaboration

Paradigm, these intervening variables validated the positive association between Speed DIAL scores and scores on the Virginia Standards of Learning third-grade tests.

DEDICATION

They say that behind every good man, there stands a better woman. In my case, there are three. This study is dedicated to three special women in my life. They are: Victoria Bolling, my granny, Evelyn Walk, my mother, and Rachel Walk, my wife. My granny was the strongest person I have ever known. She was strong in mind, body, and spirit. She instilled in me a love of reading. I can still recall all those late nights when she would nod while reading the same stories over and over again. My mother's work ethic is unparalleled. She worked her whole life, standing over a hot stove in a kitchen, to ensure that I would have a better opportunity than she had. She always found a way, even during the hard times, to make sure that I had everything I needed to succeed. I still remember vividly the first time I saw my wife as she walked into an ELPA class one cold January evening. A friend can tell you that my eyes lit up when she sat down beside me. She has a caring nature and a soft heart that encourages me to be better tomorrow than I was today. I will forever love and cherish these three women who have greatly influenced the man that I am and the man that I want to be.

ACKNOWLEDGMENTS

I would like to extend my heartfelt thanks to my father Jesse Walk for providing me with love and support throughout this experience.

I wish to thank my chairperson, Dr. Louise MacKay. Dr. MacKay always had faith in my abilities, even when I didn't. She took the time to listen, and I always felt better leaving than I did coming in.

I would also like to acknowledge my committee: Dr. Hal Knight, Dr. Elizabeth Ralston, and Dr. Terrence Tollefson. Each provided valuable input and support that helped to make this experience possible.

I must also recognize the positive influence that the late Dr. Russell West had on my doctoral experience. It is a rare person who can mold and influence while still making you feel confident in your abilities. He will be missed.

I would also like to thank my statistician, Dr. Susan Twaddle, and my editor Ms. Debby Bryan for a job well done.

Finally, I must recognize my friends in cohort Z: John, Richard, Lisa, Ginger, Pam, Gary, Morgen, and Rick . . . you all made those long evenings in class bearable . . . dare I say, enjoyable. The experiences shared and the stories told will be remembered years from now when all the theories and readings will be long forgotten. I wish you and yours all the very best in your future endeavors.

CONTENTS

Chapter Page

Chapter Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER 1

INTRODUCTION

With legislation such as *No Child Left Behind* placing greater accountability on educators, the early identification of children who may be at risk for failure is of prime importance. In the push to raise educational standards, the earlier assessment of children has translated into beliefs about what second graders, first graders, kindergartners, and ultimately what preschoolers are expected to know and be able to do (Shepard & Smith, 1986). Goal One of the 1989 Educational Goals states that, "By the year 2000, all children in America will start school ready to learn" (*Action Team on School Readiness*, 1992). Clearly, five years after the deadline, we are still working toward that goal. However, we must recognize that there is increased awareness of the importance of early childhood education. Along with accountability has come the push to provide earlier remediation. Middle-school classrooms filled with underperforming students struggling to catch up after years of falling further and further behind have validated the need for early intervention.

Today, children entering preschool or kindergarten are likely to take one or more screening or readiness tests as they begin their educational journey toward adulthood. According to Wenner (1995), the impetus for such screening did not emerge with this latest push toward school reform. Screening and diagnosing children at an earlier age has been on the rise for the last three decades. The push for early screening instruments was provided mostly by laws and the accompanying funding that began in the 1960s with such legislative acts as the *Maternal Child Health Act, the Mental Retardation Act, and the Educational Opportunity Act* (Paget & Nagele, 1986). *The Education of the Handicapped Amendment* and the *Education for All Handicapped Children's Act of 1975*, also known as PL 94-142, specified that every state would identify, evaluate, and provide intervention for all children who needed it (Meisels, 1976). Later,

PL 99-457, *Education of the Handicapped* passed in 1986, expanded the identification and duty to provide services further down to those students age three and higher. To meet the challenge of PL 94-142 and PL 99-457, hundreds of new screening tests would be published over the next 30 years (Southworth, Burr, & Cox, 1980).

Those in favor of screening children argue that screening is important so that children who need additional services can be identified and served. Schools often cite the need to "identify developmental lags, to see who is developmentally ready, to identify those who should stay home for a year, to have children placed in a developmental class, or to determine readiness for school" (Durkin, 1987, p. 767). A recent survey of current screening practices by Costenbadler, Rohrer, and DiFonzo (2000) found that of the school districts participating in the survey, 51% used a published screening test and 30% reported that they used locally created screening tests. Those in favor of screening pointed to research that indicates the instruments used for screening purposes do possess predictive factors. These supporters of screening note that identification can lead to preventative intervention services for these children.

Critics of preschool and kindergarten screenings often cite the confusion and misunderstandings about screening. Developmental screening is often lumped together with other tests such as readiness or intelligence tests. Many school districts fail to distinguish between readiness and screening tests and often use the two interchangeably toward the same purpose (Meisels, Wiske, & Tivnan, 1984). Detractors also often cite the limited establishment of reliability and validity as reasons to question the use of such screening instruments (Meisels, 1987a). There is the commonly held belief that by screening children, educators will be able to predict how well they will perform academically in the future (Stone & Gridley, 1991). However, as Drieling and Copeland (1988) noted, many of these screening tests lack the follow-up research to show a demonstrated empirical link between screening test performance and future outcomes.

Statement of the Problem

Many school districts throughout the United States use the DIAL, Developmental Indicators for the Assessment of Learning-Speed DIAL, as both a predictive and prescriptive screening instrument. As noted by Klein (1977) what is most needed is the establishment of sufficient predictive validity information. The purpose of this research study was to examine the predictive validity of the Speed DIAL developmental screening instrument. In this study, the Speed DIAL scores were correlated with the outcome criterion, the Virginia third-grade Standards of Learning tests that were administered nearly five years later. The primary purpose of this research was to validate psychometric integrity of the Speed DIAL by establishing its predictive validity.

Research Questions

The following research questions guided this study:

Research Question 1*:* Is there an association between Speed DIAL classification (*potential delay* versus *okay*) and pass-fail outcomes of the Virginia Standards of Learning third-grade math test scores?

Research Question 2*:* Is there an association between Speed DIAL classification (*potential delay* versus *okay*) and pass-fail outcomes of the Virginia Standards of Learning third-grade total English test scores?

Research Question 3: Is there an association between gender, age, and Speed DIAL scores?

Research Question 4: Is there an association between Speed DIAL scores and the intervening variables: retention status, special education status, and number of years of participation in Title I math, Title I reading, and PALS?

Research Question 5: Is there an association between the Virginia Standards of Learning third-grade math test scores and gender, age, Speed DIAL scores, retention status, special education status, and number of years in the Title I math program?

Research Question 6: Is there an association between the Virginia Standards of Learning third-grade total English test scores and gender, age, Speed DIAL scores, retention status, special education status, number of years in the Title I reading program, and number of years in the Phonological Awareness Literacy Screening (PALS) program?

Significance of the Study

The fact that many schools use screening tests such as the Speed DIAL Developmental Screening Test underscores the need to gather and analyze the available data about the assessments used as a predictive indicator. As noted by Costenbadler et al. in the year-2000 survey of school screening practices, 26% of the respondents (385 school districts) used a version of the DIAL assessment. Also, 52% of those in the participating districts indicated that they used the results of such screening instruments to either "refer children for further evaluation, place a child's name on a monitoring list, advise that the child delay school entry, or provide the child with early intervention programs" (Costenbadler et al., p. 328). Because such decisions have far reaching repercussions for the future academic success of a child, it is important that the instruments used, in this case, Speed DIAL, possess a high degree of reliability and validity (Gredler, 1997).

The Speed DIAL is used to identify children who need further evaluation in order to determine whether intervention services might be beneficial. Teachers and administrators may use such data to make crucial decisions concerning curriculum and instruction for individuals as well as whole groups of students.

Through the examination of this one screening instrument, Speed DIAL, this study might be the impetus for more thoughtful discussion, and more importantly, action in the area of

preschool screening. This study might impact policy at the local school level and possibly focus more attention on the importance of early childhood education.

This research might also help address shortcomings in the existing literature concerning the use of Speed DIAL as a screening instrument. Whereas researchers have studied the predictive validity of the DIAL and DIAL-R, the DIAL-3, and especially the Speed DIAL, versions of the tests have received little attention among the research community (Chew & Lang, 1990; Docherty, 1983; Vacc, Vacc, & Fogleman, 1987). This is true despite the fact that surveys have confirmed that both tests are used to screen thousands of children across the United States each year.

An examination of this study should add to the discussion surrounding early childhood programs and especially the use of screening instruments to predict future academic success. It is the hope of this researcher that such discussion leads to both answers and questions for further research.

Definitions

The following are definitions of terms used in this study:

1. *At-Risk Students*: Those students who have a higher than average chance of school failure (Stone & Gridley, 1991).
2. *Developmental Screening Tests*: A brief assessment procedure designed to identify children at risk for developmental problems or delays and who should proceed to a more intensive level of referral for assessment (Mantzicopoulos, 2000).
3. *DIAL (Developmental Indicators for the Assessment of Learning):* Developed over a quarter century ago under the supervision of the Illinois State Board of Education as a screening procedure that would identify young children with either current or potential learning problems (Czudnowski & Goldenberg, 1998).

4. *DIAL-3:* A 1997 revision of the DIAL-R, Developmental Indicators for the Assessment of Learning-Revised. A norm-referenced and individually administered developmental screening test designed to identify young children in need of more diagnostic assessment. The DIAL-R was modified to screen in the five developmental areas mandated in the *Individuals with Disabilities Education Act* (Czudnowski & Goldenberg).
5. *DIAL-R:* The first revision of the DIAL in 1983. The objective of the revision was to improve content, materials, and procedures. This revision was also standardized on a national sample. The age range for the assessment was also extended (Czudnowski & Goldenberg).
6. *Intervention*: The act of providing supplementary programs for a child identified as one who might benefit from such programs (Meisels, 1987a)
7. *Okay*: A decision that can be made about a child as a result of Speed DIAL screening. A performance that is within the average range of scores, indicating age-appropriate skill development based on a cut-off level chosen by the screener (Czudnowski & Goldenberg).
8. *Older*: For the purposes of this study, older students are those 1999-2000 prekindergarten students who were born before April 1, 1995, and the 2000-2001 prekindergarten students who were born before April 1, 1996.
9. *Phonological Awareness Literacy Screening (PALS)*: A diagnostic instrument used to assess student reading ability, including strengths and weaknesses. Can be used in kindergarten through third grade. (Phonological Awareness Literacy Screening Technical Reference, 2004).
10. *Potential Delay*: A decision that can be made about a child as a result of Speed DIAL screening. A performance that is below the average range of scores indicating a

potential learning delay based on a cut-off level chosen by the screener (Czudnowski & Goldenberg).

11. *Predictive Validity*: A screening instrument's ability to forecast future academic performance (Bredekamp & Shepard, 1989; Kurdek & Sinclair, 2001).
12. *Readiness Tests*: Measures of curriculum related skills that a child has already acquired (Meisels, 1987b). "Skills that are typically prerequisites for specific instructional programs" (Shepard, 1997, p. 92).
13. *Retention*: The requirement that a student who has been in a grade for a full year complete that grade again for the next school year due to academic or maturational reasons (Jackson, 1975).
14. *Special Education*: "Individually planned, systematically implemented, and carefully evaluated instruction to help learners with special needs to achieve the greatest possible personal self-sufficiency and success in present and future environments" (Heward, 1996, p.10).
15. *Speed DIAL*: A shorter version of the DIAL-3 that is a norm-referenced, individually administered developmental screening test designed to identify young children in need of more diagnostic assessment. It assesses motor, concepts, and language skills (Czudnowski & Goldenberg).
16. *Test Reliability*: The dependability that test results for a given individual will be similar over successive administrations (Neill & Medina, 1989).
17. *Title I Programs*: Refers to programs to schools that receive funds under Title I of the Elementary and Secondary Education Act (ESEA). Title I supports programs to improve the academic performance of students from low-income families (United States Department of Education, 2003).
18. *Validity*: In testing, the meaningfulness and usefulness of the inferences or conclusions that can be drawn from a test or test score (Messick, 1980).

19. *Younger*: For the purposes of this study, younger students are those 1999-2000 prekindergarten students who were born on or after April 1, 1995, and the 2000-2001 prekindergarten students who were born on or after April 1, 1996.

Delimitations and Limitations

This study was delimited by a number of factors.

1. This study was delimited to students enrolled in preschool programs in six public schools located in a rural southwestern Virginia school system during the 1999-2000 and 2000-2001 school years.
2. This group may not be representative of all preschool children.

The limitations include:

1. This study was limited by the accuracy and ability of educational professionals to effectively screen students using the Speed DIAL.
2. This study was limited to those preschool students who remained enrolled until third grade.
3. This study was limited by the extent to which teachers varied in overall effective teaching practice.

The main limitation of this study is one of limited generalizability. The results may only be generalized to the population being studied.

Assumptions

Several assumptions were made for this study. First, it was assumed the preschool students who were administered the Speed DIAL instrument represented a heterogeneous group of children. It is assumed that all testing instruments were administered in a standardized manner and that the instruments used in this study are valid and reliable.

Overview of the Study

This study contains five chapters. Chapter 1 provides an introduction to the study, statement of the problem, research questions, significance of the study, definitions of terms, delimitations and limitations, and assumptions. Chapter 2 presents a review of the related literature including: history of preschool screening, related theories concerning such screening, the difference between developmental screening tests and school readiness tests, concerns related to developmental screening tests, review of the DIAL-3 assessment, review of Speed DIAL assessment, the origin of DIAL assessment, Virginia Standards of Learning Tests, DIAL-3 and Speed DIAL validity and reliability, related DIAL studies, and related issues in screening instrument predictive validity. Chapter 3 includes research methodology and design. Chapter 4 includes a discussion of the results and Chapter 5 presents the summary, conclusions, and recommendations.

CHAPTER 2

REVIEW OF THE LITERATURE

There are 11 sections in this review of the literature. The first four sections highlight issues dealing with the practice of screening children prior to entrance into preschool and/or kindergarten. Included in these first sections are: history of intelligence testing, child development theories related to preschool screening, and misconceptions and concerns about the use of screening instruments.

Sections five through nine address the DIAL screening instrument used in this research study. Both the DIAL-3 and the shorter version, Speed DIAL, are examined in terms of their origins, features, and how they are used. Special emphasis was given to DIAL-3 and Speed DIAL reported validity and reliability. Finally, this section also includes special consideration of past research studies related to the original DIAL and DIAL-R.

Section 10 details the outcome-criterion measure used in this research study, the Virginia Standards of Learning third-grade tests. Validity and reliability of the Virginia Standards of Learning third-grade test were also a focus of this section. Finally, in the last section of the literature review, the topic of predictive validity was introduced and explained.

There exists an abundance of articles and studies on numerous preschool screening instruments. All of the research relating to the DIAL instrument was limited to the original DIAL and its predecessor the DIAL-R. This researcher was unable to locate any research studies related to the most recent versions, DIAL-3 or Speed DIAL. Many of the articles and studies are dated because they related to the older versions of the instrument. The limited available research establishes the impetus for conducting this research study.

History of Intelligence Testing

Through ages, humans have pondered and tested ideas about the nature of intelligence. As societies further developed in the 19th century and greater emphasis was given to formal education for the masses, such discussions about intelligence expanded (Paget & Bracken, 1983). With the emergence of formal educational systems, researchers of the time went about the task of developing structures to distinguish those children who might gain from a formal education and those who might not benefit (Goodenough, 1949).

A French doctor, Seguin, was one of the first physicians in Europe to suggest that those individuals with mental insufficiencies could be improved with training (Goodenough, 1949). According to Goodenough, Seguin organized his own school in 1837 to specifically train mentally challenged individuals so they might be better integrated into society. Because of Seguin's research of special needs individuals, he had a positive effect on the movement to establish early assessment instruments and practices throughout the rest of France and America (Goodenough).

According to Paget and Bracken (1983), Cattell, a German scientist who later immigrated to the United States, was the first to use the term "mental tests" in an 1890 article. With the influence of Cattell and Seguin, educators and scientists in the United States began to develop instruments to measure human cognitive development. These early tests were often poorly created with issues of validity, controlled conditions, sampling techniques, and reliability (Paget and Bracken).

Goodenough (1949) also reported that in 1905, a French scientist, Binet, developed a 30-item test to assess cognitive development. Until that time, most assessments were based on sensory functions. Binet disagreed with these early assessments and created an instrument that could evaluate the individual's reasoning, judgment, and comprehension as a measure of cognitive formation rather than sensory functioning (Goodenough). Over the next few years, the Binet scale, as it was referred to, went through several revisions. The scale was quickly

translated into English and scientists and educators throughout Europe and America were trained to administer the test (Goodenough). As explained by Paget and Bracken (1983):

> During the years between 1920 and 1940, considerable time and effort went into formulating answers to three major questions in regard to assessment. First, what are the characteristics of normal young children? Second, is intelligence determined by heredity or environment? Finally, what can be done to improve assessment devices designed to test the ability of young children? These questions were raised not only by the scientists and academicians of the time, but also by the public at large. (p. 6)

Five years after the creation of the Binet scale, the assessment of children enrolled in schools was well under way (Stott & Ball, 1965). However, children of preschool age were given limited consideration by academicians or scientists when developing assessment instruments. A few of the assessments that were expanded downward to include the preschool years often included problems with standardization, reliability, and validity (Stott & Ball).

In the 1920s and 1930s, universities across the United States organized new instructional departments devoted entirely to the study of children (Paget & Bracken, 1983). In 1937, Termer and Merrill published a revised version of the 1916 Stanford-Binet. Researchers began to do more than just study children. In addition, these same researchers began to analyze the tests. The components of what constituted human intelligence were beginning to emerge. Factors such as verbal ability, numerical ability, mechanical ability, and attention along with comprehension, word fluency, space, memory, number, and induction were the new catchwords in this emerging field of study (Paget & Bracken).

In the 1930s and 1940s, the Great Depression and onslaught of World War II brought about widespread change in the American society (Paget & Bracken, 1983). It also proved to be a time of transition for the field of child study and assessment. The introduction of child labor laws and the better enforcement of compulsory education laws ensured that more children would attend school. At the same time, the introduction of women in the workplace necessitated the creation of childcare centers across the country. This influx of children both school aged and

preschool aged children moving into quasi-controlled environments ensured that researchers of the day were not left with a shortage of students to study and assess (Paget & Bracken).

Once again, with greater numbers of students and assessment instruments, there came greater scrutiny on the assessments themselves. Many researchers performed correlation and longitudinal studies to test the predictive validity of these new assessments (Paget & Bracken, 1983).

In 1949, the Wechsler Intelligence Scale for Children was introduced. Also published were the Cattell Infant Intelligence Scale, the Northwest Infant Intelligence Scale, and the Leiter International Performance Scale (Paget & Bracken, 1983). These new instruments were far more exact than their predecessors were. There was a growing voice within the academic community supporting the idea that environment and experience greatly impacted cognitive development in ways that made quantitative assessment less than foolproof (Paget & Bracken). Paget and Bracken explained that the sensory testing models were being replaced by multi-factored designs:

> No longer could intelligence be considered a general unitary ability. It was becoming increasingly apparent that an individual's level of functioning was not dependent solely on mental activity. With the popularization of Freudian theory, psychologists and educators began considering personal and social variables as important components of overall functioning. (p. 10)

The greatest impact on preschool assessment in the 1960s was the introduction of the federal government into public education (Paget & Bracken, 1983). In 1965, the *Elementary and Secondary Education Act* ensured that the poorest and neediest among America's children would have greater access to a quality education. At the same time, the importance of early childhood education and assessment was being highlighted by psychologists, researchers, and educators alike. The result was Head Start, the federal government's first introduction into the realm of preschool funding (Paget & Bracken). Proponents of the Head Start program argued from the onset that a performance-based evaluation component should be included in the Head Start program. Along with the publicly funded Head Start program, there also came a host of

privately funded early intervention programs. Head Start programs across the nation were organized with goals and objectives individualized to meet the needs of different populations. These systematic organizations emphasized nurturing environments, academic methods, and child-centered instruction (Paget & Bracken). Although these programs varied in purpose and theory, they did not differ in program evaluation. Provisions within the legislation stipulated that funding was contingent upon measurable student outcomes. Few of the assessments at the time were created with preschool aged children in mind. Many assessments were indecisive or not designed to measure the new program goals or objectives. As a result, many new assessment instruments were developed over the next 20 years. The authors of these new assessments would place less emphasis on the measurement of intelligence and more on the measurement of achievement (Stott & Ball, 1965).

Screening children for entry into public schools was also a consequence of the medical community's performing screenings related to children's health (Potton, 1983). Such early medical screening detected problems so that a course of treatment could be rendered (Frankenburg, 1985).

The *Education for All Handicapped Children Act* of 1974 (PL 94-142) and its 1986 amendment (PL 99-457) mandated the early identification of children who exhibited characteristics that might cause them to experience learning problems in school settings. The 1986 amendment expanded the right of a free and appropriate public education to children of ages three to five before they formally began instruction in public schools (Stone & Gridley, 1991). Many states now have procedures for screening children prior to entering kindergarten.

Within the past 20 years, greater accountability had many school districts' educators looking at ways to initiate earlier intervention. Educators viewed this intervention as of great value because the prevention and treatment of learning disabilities could possibly include the prevention of behaviors and emotional problems that might influence academic performance (Lazar & Darlington, 1982).

Child Development Theories and Screening

Several different child development theories have influenced the way children are screened. Pioneers in the area of child development such as Gesell (1940) supported the maturational or developmental theory as it is sometimes called. Gesell viewed child development as a biological process that took place in stages. He theorized that child growth and development occurred in a predictable way for all children and that this development could be measured by screening tests. A child needs to be developmentally ready to enter school. For this reason, proponents of the maturational theory postulated that all children should be screened before entering school to ensure that each child is ready (Hunt, 1969). Taken a step further, some would view screening not only as an appropriate method of determining which children might be unready for school but also to identify those who might benefit from a transitional program or intervention program (DeCos, 1997).

The school readiness view relied heavily on the idea that a child's age and social growth were interrelated. Therefore, by a certain defined chronological age, a child should be able to begin school and be successful. Most school systems in the United States have set chronological points that define when a child should begin school. Effective early screening instruments could discourage the sole use of chronological age as a benchmark for school entrance that often result in problems for the child (Obrzut, Bolocofsky, Heath, & Jones 1981). Researchers such as Lichtenstein and Ireton (1984) argued that educators dramatically increased the chances of failure for a child when they forced him or her to begin school before he or she was developmentally ready.

The cognitive development theory by Piaget (1970) supported the idea that children should begin their formal education only when they were developmentally ready. According to Piaget's postulates, children go through a series of stages in their cognitive development. He recognized that experiences and environment were important to a child's development. Piaget reported that a child's particular stage of development would influence how that child would

cope and respond to environmental experiences. Cognitive theorists view effective screening as a means to pinpoint where a child is on the developmental continuum, and more importantly, use the information to plan, organize, and implement an educational program to match each child's stage of development.

Gagne's (1985) cumulative-skills theory suggested that human learning was sequential in that learning simple skills lead to the acquisition of more complex abilities. Gagne noted that new instruction should be built upon what a child has already acquired; much like the foundation upon which a house is built, so too is a child's earliest learning important to his or her later success.

Another cognitive theorist, Ausubel, defined child development as a mixture of both maturational and social experiences (Ausubel, Sullivan, & Ives 1980). Ausubel et al. theorized that nurture (experiences and environment) was equally as important as nature on a child's cognitive development growth In the area of screening children, Ausubel et al. suggested it was a mistake to only include the maturational indicators while ignoring the environment. They theorized that a lack of experiences and/or a negative learning environment were just as important as maturation for a child's cognitive development.

The constructionist and social cognitive theorists advanced the idea that children learn and develop by interacting with their parents, teachers, and others. The sociocultural theorists viewed children as active participants in their own learning, and that cognitive development was the result of these social interactions. In terms of screening, sociocultural theorists seemed to agree that the degree to which a child could positively interact with his or her teacher was of more importance than what he or she already could or could not do (Vygotsky, 1978).

Developmental Screening Tests and School Readiness Tests: There Is a Difference

As public schools' educators venture further into the area of early childhood screening, they have a greater duty to ensure that students identified by these early screenings are given the

appropriate intervention. This method of identifying and intervening has been made more problematic because of the confusion about developmental screening tests and readiness tests (Meisels, 1987b). Different goals require different tests. Meisels (1986) distinguished between the two types of tests. Developmental screening tests are used by educators to identify those students who will likely experience learning problems, whereas readiness tests inform us about a child's preparedness for benefiting from a specific program or curricula (Meisels, 1986). Developmental screening tests do give a brief look at the developmental abilities that are associated with a student's future academic success. Readiness tests measure which skills a child already possesses but give little information about a child's ability to progress to a higher level of mastery (Meisels, 1986).

Developmental Screening Test Concerns

Many advocates of early childhood education take issue with developmental screening tests because they fear the label that may be placed on a student with certain scores. Their concerns lie with the decisions that school officials will make regarding intervention based on these screening results. Some of the negative effects of intervention caused by screening scores could include: (a) moving children out of the regular classroom, (b) damaging a student's self-esteem, and (c) lowering a teacher's expectation of a student (Hobbs, 1975; Mallory & Kerns, 1988; Rosenthal & Jacobson, 1968).

Others pointed out the difficulties of testing this age group. The preschool years are a period of change. The characteristics of this age group make them less than perfect test takers. These children are immature and often lack the social and cognitive skills necessary to participate in the screening. These children often can not sit still for a measured length of time and may be unable to separate from parents long enough to complete the screening process without anxiety (Bracken 1987; Bracken 1992; Clark, 1982).

There were also problems found with using tests that have little or no established reliability and/or validity. Test reliability hinges on dependability or how often the identified results could be obtained if the test was given again (Meisels, 1987a). A test's validity is a measure of the test's accuracy. Does the test measure what it says it will measure? The students being tested are in a period of great change and the screening test used to evaluate these students must be stable and accurate (Meisels, 1987a). Bredekamp and Rosegrant (1992) were concerned that screening test results did not lead to more individualized instruction: instead, they lead to more kindergarteners being left home another year, more children retained in a grade, or more children grouped into a remedial classroom based on ability or developmental level.

Developmental Indicators for the Assessment of Learning-Third Edition (DIAL-3)

The DIAL-3-Developmental Indicators for the Assessment of Learning Third Edition is a developmental screening assessment. According to Czudnowski and Goldenberg (1998), the DIAL-3 is a norm-referenced, individually administered preschool screening test that usually takes 25 to 30 minutes to complete. It is designed to be used with children between the ages of 3.0 and 6.11. It is used to identify those with a potential to develop learning difficulties. The children are asked to complete tasks that mirror expected behaviors of students in a classroom. The DIAL-3 includes five screening areas: (a) motor, (b) concepts, (c) language, (d) self-help development, and (e) social development (Czudnowski & Goldenberg).

Speed DIAL Preschool Screening Assessment

The Speed DIAL is a revised and shorter version of the full DIAL-3 preschool screening assessment. The Speed DIAL preschool screening instrument includes 10 DIAL-3 items and usually takes half the time to administer as compared to the full DIAL-3. Within the Speed DIAL test, 9 of the 10 tasks fall within the 48 month age span. The 10 tasks pulled from the DIAL-3 for the Speed DIAL were chosen as most representative of the content of the full DIAL-

3 and most predictive of possible learning problems based on research. According to Czudnowski and Goldenberg (1998), the 10 tasks include:

1. demonstrating jumping, hopping, and skipping abilities;
2. identifying body parts like nose, knee, or thumb;
3. building a structure using three blocks;
4. copying symbols such as plus, square, triangle, and letters;
5. recognizing colors;
6. naming colors rapidly;
7. identifying concepts on a picture such as longest, most, least, and biggest;
8. identifying actions on a picture such as fly it, drive it, tell time;
9. demonstrating knowledge of letters and sounds; and
10. demonstrating problem-solving skills.

The Speed DIAL produces a total score, that, according to Czudnowski and Goldenberg, "yields solid reliability and high correlation to the full DIAL-3 score" (p. 2).

As noted by Czudnowski and Goldenberg (1998), the DIAL-3 and Speed DIAL are not readiness tests. Both are developmental screening instruments that view development as a range of functioning and not a fixed point. The DIAL-3 and Speed DIAL are meant to provide a meaningful indicator for identifying children who may have learning problems and need specific intervention.

The DIAL-3 contains three subtests that are designed to measure motor, concept, and language development. "A child is identified as *potential delay*, if two of the subtest category scores fall below -1.5 standard deviations of the respective subtests. According to Suen, Czudnowski, and Goldenberg (1998), "Children who score within the range of the respective subtests on two of the three subtests are identified as *okay*" (p. 674). Czudnowski and Goldenberg wrote:

> DIAL-3 and Speed DIAL are designed to identify both children *at-risk* and developmentally delayed, keeping in mind the importance of the succeeding steps to

> diagnostically verify the degree of delay and the impact of the delay on future developmental growth as well as academic standing. (p. 6)

Czudnowski and Goldenberg also identified two basic applications of DIAL-3 and Speed DIAL for public schools as:

1. identifying children with possible learning difficulties who may benefit from further testing, intervention, or special services; and
2. identifying children who may be *at-risk* and who may benefit from intervention programs provided by the school. (p. 7)

The authors, Czudnowski and Goldenberg, also pointed out:

> DIAL-3 and Speed DIAL are neither readiness tests, intelligence tests, nor diagnostic tests. They do not measure innate abilities or identify those children with brain dysfunction. DIAL-3 and Speed DIAL are simply first steps to identify young children at the lower end of the continuum of developmental skills who may be in need of additional services, thus they must be followed by a more complete assessment before developing special programs outside the regular classroom. (p. 7)

The Origin of Developmental Indicators for the Assessment of Learning (DIAL)

Over three decades ago, the Illinois State government finalized legislation that called for the creation of a screening instrument that could identify children of preschool age who may have learning problems (Czudnowski & Goldenberg, 1998). The job was given to the Illinois State Board of Education to develop a screening test that met the criteria called for in the legislation; the result was the creation of the Developmental Indicators for the Assessment of Learning (DIAL) screening test (Czudnowski & Goldenberg). The DIAL was designed to meet 10 criteria established by an advisory board. The authors noted that while other tests at that time met some of these criteria, no test had met all 10. As noted in the instrument manual, the criteria specified that the test should:

1. be a screening test rather than a diagnostic test, relatively short, of a surface nature, and designed to indicate the possibility of a variance in development;

2. cover the age range of 2.5 to 5.5 years to assist local school districts in identifying preschool children requiring special education services;
3. be administered individually (the only way to assess gross motor skills, articulation, and expressive language adequately), but in a group setting that simulates a typical preschool or kindergarten classroom;
4. take about 30 minutes to fit the young child's short attention span and need for organizational systems to screen many children;
5. cover many areas of development (be multi-dimensional);
6. be noncategorical and attempt to identify *at-risk* children regardless of the reason for the potential learning problem;
7. be scored on the basis of observable performance rather than the subjective opinion of a tester;
8. be process oriented as well as product oriented;
9. be applicable to culturally diverse populations; and
10. be normed on a large, stratified sample. (p. 11)

Validity and Reliability of the DIAL-3 and Speed DIAL

Internal reliability relates to the consistency of a test taker's responses to items in a test. This internal consistency focuses on the degree to which the individual items are correlated with each other. Cronback's coefficient alpha is used to establish internal reliability for the DIAL-3 and Speed DIAL screening instruments (Czudnowski & Goldenberg, 1998). As pointed out by Czudnowski and Goldenberg, the DIAL-3 and Speed DIAL manual noted:

> Coefficient alpha yields a result that is the equivalent of the mean of all possible split-half reliability coefficients. To compute a split-half reliability, the assessment is divided into halves. This split can be achieved in different ways. One way is separating the test into odd and even numbered items. Another possible split can be obtained by taking the first 25% of questions and grouping them with the last 25% of questions, leaving the middle 50%. (p. 79)

According to Salvia and Ysseldyke (1998), “Researchers generally accept a standard of .80 when measuring the reliability of screening tests” (p. 163). The DIAL-3 total score has a reliability score of .87. Czudnowski and Goldenberg also noted that Speed DIAL has an alpha reliability score of .80. According to these authors, the Speed DIAL, being a shortened form of DIAL-3, would be expected to have slightly lower reliability. However, "its medium reliability of .80 is adequate for a brief developmental screening measure” (p. 80).

A test-retest reliability measure is established by administering the same test twice and comparing the scores. Consistency across time was established for the DIAL-3 and Speed DIAL through test and retest methods (Czudnowski & Goldenberg, 1998). According to Czudnowski and Goldenberg, the DIAL manual's authors:

> The DIAL-3 was administered twice to 158 children drawn from the standardized sample. The DIAL-3 total (.88 and .84 for the two groups) and the Speed DIAL (.84 and .82) have very satisfactory test-retest reliabilities that are above the .80 criterion (p. 80).

The validity of a screening instrument is the degree to which it measures what it is supposed to measure. The developers of DIAL-3 and Speed DIAL used several means to establish validity. Content and criterion-related validity evidence was confirmed by comparing results of the DIAL-3 with other criteria such as: diagnostic evaluations, success in the classroom, or placement in special intervention programs. The creators of DIAL-3 and Speed DIAL also included a thorough description of test content, a theoretical rationale for the screening, and various sampling of test construction and administration along with constant data analysis to validate the entire DIAL-3 and Speed DIAL screening instruments (Czudnowski & Goldenberg, 1998).

Validity was also accomplished by correlating the DIAL-3 screening instrument with its previous versions such as DIAL-R. To establish this connection, the researchers performed a study in which groups of students were given both versions of the screening instrument to determine whether there were mean differences in scores between the two versions of the test (Czudnowski & Goldenberg, 1998). Both the DIAL-3 and Speed DIAL exhibited high

correlation (.91 and .87 for younger samples, and .84. and .81 for older samples). According to Czudnowski and Goldenberg, "These results provide good support for the convergent and discriminate validity of DIAL-3 scores. The total means scores show a difference of one to two points. The difference between the Speed DIAL and DIAL-R total is even smaller" (p. 83). The researchers also used intercorrelation within different parts of the DIAL-3 itself to provide validity. The correlation between DIAL-3 total score and the Speed DIAL score was .94, which suggested the use of Speed DIAL as a valid, alternative-screening instrument (Czudnowski & Goldenberg).

Finally, the validity of the DIAL-3 screening instrument was judged in terms of its correlation with other widely-used preschool screening instruments and other diagnostic assessments. In the area of preschool screening test, the DIAL-3 was compared with the Early Screening Profiles Test (ESP), the Batelle Developmental Inventory Screening Test (BDIST), the Bracken Basic Concept Scale Screening Test (BBCS), and the Brigance Preschool Screen. Correlation among area scores with the DIAL-3 and other screening instruments were in the moderate range. However, with some instruments such as the Batelle Screening Test, higher correlations existed between the total scores. It should also be noted that when comparing these screening instruments with the DIAL-3, the researchers also compared the total Speed DIAL score to ensure the validity of this brief, alternative assessment (Czudnowski & Goldenberg, 1998).

Looking beyond the total scores and examining comparisons between DIAL-3 concepts and subtests within the other instruments, moderate to high correlations could be found (Czudnowski & Goldenberg, 1998). Some tests such as the Bracken Screening only test basic concepts whereas others such as the Peabody Picture Vocabulary only test receptive vocabulary. Looking at the correlations within the tests, the DIAL-3 total scores compared moderately with ESP language subscores with a correlation of .63. The Brigance Preschool Screen compared high (.79) with the DIAL Language Area. Likewise, the DIAL-3 language had a high correlation

of .76 with the Differential Ability Scales Verbal Cluster (Czudnowski & Goldenberg). Absolute correlations between the DIAL-3 and other scoring instruments were impossible because of design differences. However, when one looks within the features of each assessment, a significant correlation among concepts and areas demonstrates strong association with the construct validity of the DIAL-3 and Speed DIAL Preschool Screening Test (Czudnowski & Goldenberg).

Related Developmental Indicators for Assessment of Learning (DIAL) Studies

While internal research by the developers of DIAL-3 and Speed DIAL seems to indicate that the screening instruments are a valid and reliable measure of potential learning problems, only a handful of other researchers have tested the instrument (Vacc et al., 1987). Obrzut et al. (1981) researched the validity of the DIAL and determined that only the DIAL Communication area was a good predictor of future achievement based on its correlation with the Metropolitan Readiness test, and that the Concepts subtest was the only significant predictor of how a student would perform on the kindergarten progress report. Docherty (1983) conducted a study and found that the number of children classified by the DIAL as needing further evaluation (7%) was reasonable. Docherty also concluded that the DIAL subtests seemed to exhibit moderate intercorrelations with a range of .34 to .63. Docherty did raise concerns about DIAL norm validity because of consistent gender differences that were partial to females. However, Docherty explained the 1.1-point difference of the scales to be related to the large population available. Overall, the study's author found the experimental results to be generally positive of the DIAL.

While these researchers looked at the correlations between DIAL and future performance in kindergarten, the research by Vacc et al. (1987) was the first recognized study to address the correlation between DIAL and students' performance on the California Achievement Test (CAT) for first graders. Through correlational analysis, the researchers set out to test relationships

between the DIAL subtests and the CAT subtests. The analysis provided a correlational coefficient of .67, which, according to the researchers, "indicates that the DIAL is a valid predictor of successful school performance as measured by the CAT test" (Vacc et al., p. 48). Overall, the researchers found that the DIAL Concepts subtest was the best predictor of performance as measured on the CAT. This study supported the findings of Obrzut et al. (1981).

Virginia's Standards of Learning Assessments

The Virginia Standards of Learning Assessments are comprised of multiple-choice items and writing prompts designed to assess the content of Virginia's Standards of Learning (SOL). Students in grades three through eight and high school are tested. The assessments are cumulative at the elementary and middle school level. For example, the content area of the grade- three tests contains items that cover Standards of Learning Content from grades kindergarten through three. The Standards of Learning Assessments are criterion-referenced tests. Unlike norm-referenced tests that compare a student's score with that of other students taking the test, the Standards of Learning assessments are criterion-referenced that relate each student's performance to a predetermined set of criteria (Virginia Department of Education, 2000).

Validity and Reliability

The most important criterion for establishing the validity of any assessment is whether the test truly measures what it is supposed to measure (Virginia Department of Education, 1999). In the case of the Virginia Standards of Learning tests, do the questions on the test measure the content of the Standards of Learning? The Virginia Standards of Learning Content Review Committee comprised of educators, the testing contractor, Harcourt-Brace Educational Measurement, and the Virginia Department of Education worked methodically to ensure that every test question matched the corresponding Standards of Learning. Validity was also

established by making comparisons among different assessment instruments. According to the Virginia Standards of Learning Assessments Validity and Reliability Information:

> In the content areas and grade levels where there were reasonable matches of content, school pass rates on the Standards of Learning test have been statistically correlated with national percentile ranks on the Stanford 9 Achievement Test and/or pass rates on the Literacy Passport Test. These data show a strong relationship between the relative standing of Virginia's schools on the Standards of Learning tests and both the Stanford 9 Achievement Test and the Literacy Passport Test. Though varying among grades and content areas, schools that scored well on the Stanford 9 Achievement Test or Literacy Passport Test generally scored well on related Standards of Learning tests and vice versa. (Virginia Department of Education, 1999, p. 8)

The Spearman Rank Order correlation coefficients for pass rates on the grade three Standards of Learning tests compared moderately to high with national percentile ranks on the grade three Stanford 9 Achievement Test. The correlation coefficients ranged from .67 to .78. (Virginia Department of Education, 1999)

Reliability is the extent to which a test consistently measures what it is supposed to measure. It is also the degree to which the scores are dependable. As noted in the Virginia Standards of Learning Assessments Validity and Reliability Information (Virginia Department of Education, 1999):

> A high degree of reliability within the Standards of Learning Assessments is critically important. When developing the Virginia Standards of Learning tests the developers utilized the Kruder-Richardson 20, or KR-20, as the statistical measure of test reliability for all Standards of Learning tests except writing where person separation reliability was used. The Kruder-Richardson 20 is a traditional measure designed to calculate the degree to which the test questions consistently measure the same body of content and skills. KR-20 values range from 0 to .99. Test developers aim for a test KR-20 value to be as high as possible. The KR values for the Virginia Standards of Learning tests range from a low of .80 to high of .92. (p. 19)

Related Issues in Screening Instrument Predictive Validity

There is disagreement concerning labeling of children as *at-risk* or *developmentally delayed* (Gredler, 1997). A screening instrument is considered valid if it is able to differentiate between children who may have potential learning difficulties and those who do not have

potential learning difficulties (Lichenstein & Ireton, 1984). The difficulty lies in the different expectations of abilities and the different views as to what constitutes school success. Schools are attempting to measure a wide variety of variables and different screening instruments yield different validity correlations with later criterion assessment scores. Therefore, one school system may identify a student as *at-risk* and that same student may not be identified as such in another system (Gredler, 1997). In addition, no researcher can predict the quality of instruction that a child will receive in the classroom. Educators are left with a predictive validity model where the results of a screening instrument are compared to a later criterion, such as an achievement test, and attempts are made to correlate the two in the hopes of justifying the first instrument (Gredler, 1976, 1997).

The analysis of data from the screening instrument and later criterion measures can lead to four possible outcomes. The first is a valid-positive outcome. Here, children who perform poorly on the first screening also perform poorly on the criterion test. The second outcome, false negatives, are children identified as not at-risk on the screening test but who later perform poorly on the criterion measurement. The third outcome is false positive. Here, the children are identified as at-risk by their performance on the earlier screening but perform adequately on the criterion test. The fourth and final outcomes are classified as valid negatives. These are children who are identified by the screening instrument as not at-risk and later score adequately on the criterion assessment (Gredler, 1997). Gredler (1992) identified two indices that are most important in establishing the predictive validity of a screening instrument. The first was sensitivity index, which gives the percentage of children that scored poorly on the criterion test who were identified as at-risk on the earlier screening test. The second index is the Index of Specificity. This index computes the percentage of children who were correctly identified by the original screening test (Gredler, 1992).

There have been various studies that have probed the association between early screening instruments and later academic success. Kaplan (1996) found that scores on Wechsler Preschool

and Primary Scale of Intelligence-Revised (WPPSI-R) were predictors of later academic success on the Comprehensive Testing Program (CTP). Specifically, the analysis of the correlations between WPPSI-R subtests and academic achievement areas on CTP showed that verbal subtests of the WPPSI-R were consistently correlated with achievement (Kaplan). Dziuban and Mealor (1982), in their study of 224 third graders in Florida, found a moderately positive correlation between the developmental indices on the Yellow Brick Road Screening Instrument and subsequent achievement of the Comprehensive Test of Basic Skills. The association between the two tests deserved further consideration and strengthened the idea that consideration of a child's developmental status before kindergarten was beneficial in predicting future academic performance. Horn and Packard's (1985) meta-analysis of 58 studies suggested that the screening variables with the most predictive validity were those that measured the child's ability to focus and remain focused without being distracted. They also noted that language skills and general intelligence functioning were variables exhibiting the most predictive validity. The analysis of the data also proved that low scores in the area of motor skills often cited as predictors of future performance were not as effective as were early predictors of future school success.

DeHirsch, Jansky, and Langford (1966) analyzed the results of 13 screening tests using a screening battery approach to effectively predict which children were retained in second grade. Eaves, Kendell, and Crichton (1974) used the same approach as DeHirsch et al. and reported correlations in the range of .95. Book (1980) found that a significant correlation existed between the kindergarten group designations and subsequent group test achievement in grades one through four. The research found that students' academic performance changed little over time. In this particular study, the screening test consistently predicted group academic performance through the fourth grade. More than proving the predictive validity of screening instruments, this study highlighted the importance of carefully planned intervention programs. Steinbauer and Heller (1978) reported that the receptive language indicators on a screening instrument

administered during kindergarten had a significant correlation with reading comprehension, math, and spelling on the third-grade achievement test.

Some researchers have pointed to the significance of documenting a child's development as a valid predictor. However, Gallerani, O'Regan, and Reinhertz (1982) noted that a screening instrument that measured cognition was a better predictor of future ability than an individual's developmental history. Schmidt and Perino (1985) examined a kindergarten-screening test where the dependent variables in the study included achievement scores on the Metropolitan Achievement Test, scores on the Otis Lennon School Ability Test, and students' placement in special education programs. The researchers found that a combination of subtests accurately predicted 77% of the children who would later be served by special education programs. The analysis also concluded that the same kindergarten screening could adequately identify 73% of those students who would be classified as high achievers by the second-grade assessments (Schmidt & Perino).

Agostin and Bain (1997) researched the predictive validity of the kindergarten-screening instrument and reported that both developmental skills and social skills were key predictive components. Within developmental skills, they noted factors such as self-control, cooperation, and receptive language ability to be the most important variables for the prediction and discrimination for children who might be at risk of failing or being included in intervention programs.

In a study by Chew and Lang (1990), the authors focused on the association between two kindergarten screens, The Lollipop Test and the DIAL. The dependent variables were the California Achievement Test, the Georgia Criterion-Referenced Test, and teacher reported grades. The study found a high degree of correlation between the Lollipop and DIAL screens as both displayed an equal ability to predict students' test scores and grades. Correlations between the predictive screens and the dependent variables ranged from .68 to .83.

Funk, Sturner, and Green (1986) examined the predictive validity of the McCarthy Scales of Children's Abilities (MSCA). The study highlighted the correlation among preschool scores on the McCarthy Scales of Children's Abilities and the California Achievement Test given to second graders three years later. The preschool McCarthy Scales of Children's Abilities scores displayed moderate to high correlation with the outcome, the California Achievement Test. More specifically, the researchers noted:

> Children who scored low, less than 68 on the McCarthy Scales of Children's Abilities General Cognitive Index, had significantly lower California Achievement Test scores in all grades than did children scoring within the normal range of greater than 84 on the General Cognitive Index. Seventy-four percent of those identified as scoring low on the General Cognitive Index had either failed a grade, been placed in special education classes, or were scoring in the bottom 25% in their cohort on the California Achievement Test in that year. (p. 181)

Not all studies have focused strictly on the predictive validity of preschool and kindergarten screening instruments. Simmer and Barnes (1991) conducted a study to determine the correlation between marks in first grade and later school success. Using the permanent records of 200 high school students, their grades were examined from 1st grade through 12th grade. The researchers reported a high correlation between the grades a student made in language arts and math in first grade and the grades made in high school. The correlations decreased as the students progressed; however correlation coefficients greater than .50 were discovered as late as early high school. To make a connection to early screening, the authors of the study also noted that if screening scores had been used to create meaningful instructional programs for individual students, many of the students identified as at-risk by the screening instruments might have been effectively remediated. Unlike several studies revealing the predictive validity of screening instruments, Thurlow, O'Sullivan, and Ysseldyke (1986) concluded in their research that screening instruments rarely test what they are supposed to measure. They also wrote that many of the studies on predictive validity failed to incorporate sufficient sampling. Shepard and Smith (1986) cautioned the use of screening data for anything other than teacher planning. They wrote:

> There is one overriding rule for determining test validity: validity depends on how a test is used. In the case of school screening measures, this means that some tests might be perfectly good for teachers to use in making day-to-day decisions, but would not be good enough (technically or in a court of law) to be used to place a child in a special school program. The more crucial the decision for an individual child, the greater are demands for test validity evidence and due process. (p. 83)

Within the study of test predictive validity, there seems to be confusion as to what constitutes a positive correlation. A correlation may be noted as high in one study and moderate in another. In one study. a correlation of .35 could be viewed as moderate whereas another study might report that same correlation as low. Even with studies of the same screening instrument, there is a lack of agreement. For example, Ames and Ilg (1964) examined the Gesell School Readiness Test and reported a predictive validity correlation in the moderate range of .74. Other researchers such as Popovics (1982) and Banerji (1992) who also analyzed the Gesell School Readiness Test achieved correlations ranging from .11 to .39. Such inconsistencies within the same instrument call into question the effectiveness of that particular screening instrument to predict future academic performance.

Summary

This chapter began with a brief history of preschool screening practices. Included was a review of issues and concerns associated with screening instruments. Next, a comprehensive analysis of the predictor and outcome variables for this study was presented. Following an indepth review of the related literature, the topic of predictive validity was explored. The purposes of this chapter were to establish a background and present the basis and justification for conducting this research study.

CHAPTER 3

METHODOLOGY

This chapter describes the methodology and procedures used in this research study to determine the predictive validity of the Speed DIAL screening instrument. This chapter is organized into the following sections: research design, participants and data collection, instrumentation, data analysis, hypotheses, and a summary.

Research Design

The purpose of this study was to explore the associations of Speed DIAL screening scores of preschoolers with further academic achievement as measured by scores on the Virginia Standards of Learning third-grade math, reading, and writing tests. The goal was to establish the predictive validity of the Speed DIAL screening instrument.

It was an association study designed to analyze the strength of association between variables (Gall, Borg, & Gall, 1996). This research study was nonexperimental in design because the variables were not directly manipulated. The greatest limitation of this type of nonexperimental research design is that it does not imply whether a causal association exists between the variables. A causal association exists when it is implied that one thing *causes* another thing to take place. Because correlational research studies are nonexperimental, it is impossible to factor out other variables that may influence the results (Gall et al.).

A strength of the correlational method is the ability to measure the association between variables without having to influence the particular variables. Correlational studies are widely used to study research problems in the social sciences and education fields because the correlations can be used as a basis for predictions. Correlational research provides a statistical method that allows researchers to study associations among large numbers of variables. The correlational research method gives information about the degree (low, medium, moderate, and

high) of association among variables. A correlational coefficient is the numerical strength of that association between two variables. That strength can vary positively and negatively. Positively, the strength of association can range from .00 to 1.00. A correlation coefficient of .00 indicates absolutely no association between variables while a correlation coefficient of 1.00 indicates a perfect correlation between variables (Gall et al., 1996). Depending on which study read, there seems to be some disagreement among the research community as to what constitutes a low, medium, moderate, or high correlation. In their textbook on applied statistics, Hinkle, Wiersma, and Jurs (1998) noted:

> A coefficient of .00 to .30 is considered to show little to no relationship; a coefficient between .30 and .50 indicates low correlation; between .50 and .70 moderate correlation; .70 and .90 high correlation, and a coefficient of .90 to 1.00 is considered to be a very high correlation. (p. 48)

Participants and Data Collection

The students in this study were selected because they participated in nine preschool classes and had been administered the Speed DIAL screening assessment. The population included preschoolers during the 1999-2000 and 2000-2001 school years. For both years, the combined group totaled 288 students. Of the 288 original preschoolers screened, only 205 had maintained enrollment within the school division through the third grade and consequently were given the outcome criteria, the Virginia Standards of Learning Tests. The data for this study came from the Title I and elementary education supervisors at the central office of the school system. Both supervisors served as liaisons for the study. A coding system was implemented to make certain no identifying information was used in this study. Using a spreadsheet, the Title I supervisor assigned each 1999 and 2000 preschooler a number. The Speed DIAL score, gender, and date of birth were added for each preschooler. Next, this spreadsheet document was given to the elementary education supervisor who added each preschooler's third-grade Virginia Standards of Learning test score for math and total English, if available. Finally, the elementary supervisor eliminated the student names and gave the data to the researcher.

The preschoolers attended seven elementary schools in a predominately white, lower to middle class, rural, mountainous community in the heart of the Appalachia Mountains. Each preschooler was screened using the Speed DIAL screening instrument. The population was comprised of 205 subjects ranging in age from 42 months to 52 months at the time of initial screening. There were 93 females and 112 males included in this study.

Instrumentation

Speed DIAL

For this study, the predictor variable was the Speed DIAL developmental screening test. The outcome variable for this study was academic achievement. For purposes of this study, academic achievement is defined as the student's performance on the Virginia Standards of Learning third-grade tests in math, total English, reading, and writing.

The Speed DIAL (Czudnowski & Goldenberg, 1998) is an updated shortened version of the DIAL-3. The Speed DIAL is a norm-referenced, individually administered preschool screening instrument that requires between 15 and 20 minutes to administer. It was designed to be used with children in the age range of 3-0 through 6-11. The Speed DIAL format consists of 10 DIAL-3 items selected from the three main DIAL-3 areas of motor, concepts, and language. Within the motor area, the instrument measures a child's ability to jump, hop, and skip as well as to match displays and build block designs. Within the concept area, the Speed DIAL measures a child's ability to identify colors, body parts, and concepts such as more, less, empty, and full. Within the language area, the Speed DIAL test measures a child's ability to recite the alphabet song, solve coping problems, and describe a pictorial representation. Because of its conciseness, the Speed DIAL yields only a total score.

Acceptable levels of concurrent and predictive validity are documented for the Speed DIAL. There is also available evidence of test-retest reliability and internal consistency.

The DIAL-3 was standardized between 1995 and 1997 on a nationwide sample of 1,125 children. Consideration was given to gender, race, geographic location, size of community, and chronological age. The correlation coefficient between the DIAL-3 total and Speed DIAL is reported as .94 (Czudnowski & Goldenberg, 1998.)

Before administering the Speed DIAL, screening coordinators must decide upon a particular cut-off level that is appropriate for their particular community. Each locale has its own particular demographic indicators. Some places have smaller numbers of children served by special services whereas others may provide service to greater numbers. As noted by Czudnowski and Goldenberg (1998) in the DIAL-3 manual, "Local program-wide decisions always need to be made about the most appropriate cut-off level to use, based upon the proportion of children in the community who are expected to need further evaluation of potential developmental delay" (p. 53). Once a cut-off level has been selected, results from the Speed DIAL can be interpreted and a formal designation of *potential delay* or *okay* can be assigned to each test taker. Scaled scores on the Speed DIAL range from a minimum of 0 to a maximum score of 39. The Speed DIAL scaled score may also be converted to a percentile rank.

Virginia's Standards of Learning Third-Grade Tests

The Virginia Standards of Learning third-grade tests are criterion-referenced assessments given near the end of the third-grade year. Students are tested in five areas: math, reading writing, science, and social studies/history. Students receive four scores. The reading and writing scores are combined into one total English score. For the purposes of this study, only the math and total English scores were analyzed (Virginia Department of Education, 2000).

Cut-off scores establish three performance categories for the Virginia Standards of Learning third-grade tests:

1. a score of less than 400 *does not meet standards* and indicates failing;

2. a score of 400 or greater, but less than 500 represents *proficient attainment of standards* and indicates passing; and
3. a score of 500 or greater represents *advanced attainment of standards* and indicates passing.

Raw scores on the Virginia Standards of Learning Tests are transformed into scaled scores. A raw score of zero corresponds to a scale score of zero, and a scale score of 600 corresponds to a perfect raw score (Virginia Department of Education, 2000).

Evidence confirms both the content and correlational validity of the Virginia Standards of Learning third-grade tests. Test developers used the Kuder-Richardson Formula or (KR-20) as a positive indicator of test reliability. The KR-20 reliability coefficients were .90 for the third-grade English Virginia Standards of Learning Tests and .91 for the math Virginia Standards of Learning Tests (Virginia Department of Education, 1999).

Interventions

Nearly four and a half years passed between the time the students were screened using the Speed DIAL until they were assessed using the Virginia Standards of Learning third- grade tests. Obviously, within those four years, many things happened in the lives of the students both in school and at home. Certainly, many different variables that would have affected each student's growth and performance occurred in those four and a half years. It would be impossible to document all of the day-to-day accommodations and interventions that were given to these students each and every day. However, it was possible to collect archival data from formal programs provided by the school system used for this study. Like the Speed DIAL scores and the Virginia Standards of Learning third-grade test scores, the data were provided to the researcher by liaisons within the school district. In addition to the Title I director and elementary education supervisor, the special education director also provided information. The intervention data were handled exactly in the same manner as the Speed DIAL and Virginia Standards of

Learning test scores. The same coding system was used with intervention status added to the testing variable information being provided. The documented interventions included grade retention status, special education status, years of participation in Title I math program, years of participation in Title I reading program, and years of participation in the PALS phonological awareness program.

For the purpose of this study, special education status was only documented if the student had received services under a specific learning disability or developmental delay status. Other special education services such as gifted, speech, or physical therapy assistance were excluded from this study. For purposes of confidentiality, only the question of whether a student had been served was used in this study. The number of years served was not made available to the researcher. Many school districts choose to monitor students for special education services only in kindergarten; however, in this particular school district, special education services were available in kindergarten through the 12th grade.

Each of the six elementary schools used in this study qualified for services paid for under the federal government's Title I program. The school system had a centralized Title I department with a director who supervised the Title I programs at each school. Within each school, there were teachers and aides who worked within the Title I program. Money for each school's program was based on the number of students who qualified for free and reduced-price lunch. These funds paid for Title I personnel in each school as well as materials and supplies used by the program. Each school served students in the areas of math and reading. The programs were organized both inclusively where Title I teachers went into the classrooms to assist selected students and exclusively whereby certain students were pulled from the classroom to work with Title I teachers and aides in other areas. However, one of the seven schools was operated as a school-wide Title I program where Title I teachers and aides could work with any student within the building. In the other six schools, only those students identified by the program could be served by Title I personnel. Students were chosen based on their past academic

performance. Teachers' surveys were most often used to determine which students would be served. Teachers would rank a student's ability in certain academic areas using formal and informal assessments. Using these surveys, Title I personnel would develop a list of students based on need to be served by Title I math and reading programs in each school. The Title I programs in these seven schools served children in grades kindergarten through the third grade. For this study, the number of years a student was served in Title I math and Title I reading was documented.

The school district used for this study also provides a literacy program, PALS, to students who qualify in grades kindergarten through third grade. The program's name comes from the assessment used to determine which students are served by the program. The PALS or Phonological Awareness Literacy Assessment was developed by the University of Virginia's Curry School of Education. The screening instrument is used in grades kindergarten through third grade. Teachers use the information from the PALS to decide which students may need additional help in the area of literacy development. The PALS program in this school district is available in the seven schools used in this study. In each school, trained literacy instructors work with the identified students in the areas of phonological awareness, alphabet knowledge, knowledge of letter sounds, spelling, concept of word, word recognition in isolation, and oral passage reading. PALS' scores are used to select those students with the most need in the area of literacy development. The PALS screening instrument is administered twice in the school year, both in the fall and spring. PALS program students are pulled out of the classroom setting to work in groups with the PALS instructors. This pullout time is usually no more than 30 minutes to an hour each day. Those students deemed as needing services who were not being served by other programs such as Title I or special education were given first priority in the PALS program. For this study, the number of years a student was served in the PALS program was documented.

Data Analysis

Research Questions and Hypotheses

The following strategies were used to answer the stated research questions:

Research Question 1: Is there an association between Speed DIAL classification (*potential delay* versus *okay*) and pass-fail outcomes of the Virginia Standards of Learning third-grade math test scores?

Ho1: There is no association between Speed DIAL classification and outcome on the Virginia Standards of Learning third-grade math test scores.

Research Question 2: Is there an association between Speed DIAL classification (*potential delay* versus *okay*) and pass-fail outcomes of the Virginia Standards of Learning third-grade total English test scores?

Ho2: There is no association between Speed DIAL classification and outcome on the Virginia Standards of Learning third-grade total English test scores.

Chi-square (X^2) was used to analyze these two research questions as well as the descriptive statistics for a crosstabulated table. Table 1 shows the cells for calculating the overall effectiveness, index of sensitivity, and index of specificity of the Speed DIAL classifications and SOL outcomes.

Table 1

Cell Definitions for Calculating Overall Effectiveness and Indices of Sensitivity and Specificity

Outcome	Potential Delay	Okay	Totals
Fail	Cell A	Cell B	Total Fail
Pass	Cell C	Cell D	Total Pass
Totals	Total Potential Delay	Total Okay	Grand Total

Cell, row, column, and grand total counts were used to calculate percentages for overall effectiveness and indices of sensitivity and specificity.

Some researchers argued that the use of a correlation coefficient alone provides few conclusions of a screening instrument's true predictive validity and that several indices combined can be more efficient. Gredler (1997), Lichenstein and Ireton (1984), and Leach (1980) described an additional way of analyzing data for the use of a predictive matrix. The Speed DIAL screening test determines whether a child will be classified as potential delay or okay. On the Virginia Standards of Learning third-grade test, a student is considered failing if his or her score is below 400 and passing if his or her score is 400 or above. In the predictive matrix, those children accurately classified by the Speed DIAL are represented by cells A and D. Cell A is the number of children who were classified as potential delay and who later failed the Virginia Standards of Learning test. These are referred to as valid positives. Cell D is the number of children who were classified as okay by the Speed DIAL and who later passed the Virginia Standards of Learning third-grade test. These are referred to as valid negatives. Within the matrix, those children not accurately classified by the Speed DIAL are represented by cells B and C. Cell B is the number of children who were classified by the Speed DIAL as okay but who later failed the Virginia Standards of Learning test. These are referred to as false negatives. Cell

C is the number of children who were classified by the Speed DIAL as potential delay but who later passed the Virginia Standards of Learning third-grade test. These are referred to as the false positives. The data inside this matrix can be used to establish the Speed DIAL's predictive validity in terms of overall effectiveness, index of sensitivity, and index of specificity (Gredler, 1992). Overall effectiveness is a percentage obtained by adding the valid positives in Cell A with the valid negatives in Cell D and dividing that number by the whole group. Stated another way, overall effectiveness is the number of potential delays and okays correctly classified by the Speed DIAL divided by the number of children who were given the Speed DIAL test (Gredler, 1992). Index of sensitivity is a percentage determined by taking the valid positives in Cell A, those children classified as potential delayed by Speed DIAL who later failed the Virginia Standards of Learning tests, and dividing that number by the total number of Virginia Standards of Learning grade three test failures, Cell A plus B (Gredler, 1992). Index of specificity is a percentage determined by taking the valid negatives in Cell D, those children classified as okay by the Speed DIAL who later passed the Virginia Standards of Learning test, divided by the total number passing the Virginia Standards of Learning third-grade test, Cells C plus D (Gredler, 1992). The indices of sensitivity and specificity calculate the percentages of the number of failures and passing on the Virginia Standards of Learning third-grade test that were identified originally by the Speed DIAL. Another index of importance is the percentage of children identified as potential delayed by Speed DIAL who later failed the Virginia Standards of Learning third-grade test. This percentage is determined by taking the valid positives in Cell A, those children identified as potential delay by Speed DIAL who later failed the Virginia Standards of Learning third-grade test, divided by the numbers of potential delays identified by Speed DIAL, cells A plus C (Gredler, 1992). The last index is the percentage of children classified as okay by the Speed DIAL who later passed the Virginia Standards of Learning third-grade test. This percentage is determined by taking the valid negatives in cell D, those children classified as okay by the Speed DIAL and who later passed the Standards of Learning third-

grade test, and dividing that number by the number of okays identified by the Speed DIAL, cells B and D (Gredler, 1992).

Research Question 3: Is there an association between gender, age, and Speed DIAL scores? A two-way ANOVA was used to analyze this research question. The null hypotheses for this question were:

$Ho3_1$: There is no difference between male and female Speed DIAL scores.

$Ho3_2$: There is no difference between younger and older students' Speed DIAL scores.

$Ho3_3$: There is no significant gender by age interaction.

Research Question 4: Is there an association between Speed DIAL scores and the intervening variables: retention status, special education status, and number of years of participation in Title I math, Title I reading, and PALS?

To analyze this research question, Pearson's correlations were used. The null hypotheses were:

$Ho4_1$: There is no association between Speed DIAL scores and retention status.

$Ho4_2$: There is no association between Speed DIAL scores and special education status.

$Ho4_3$: There is no association between Speed DIAL scores and the number of years of participation in Title I math.

$Ho4_4$: There is no association between Speed DIAL scores and the number of years of participation in Title I reading.

$Ho4_5$: There is no association between Speed DIAL scores and the number of years of participation in the PALS reading program.

Research Question 5: Is there an association between the Virginia Standards of Learning third-grade math test scores and gender, age, Speed DIAL scores, retention status, special education status, and number of years in the Title I math program?

Research Question 6: Is there an association between the Virginia Standards of Learning third-grade total English test scores and gender, age, Speed DIAL scores, retention status, special

education status, number of years in the Title I reading program, and number of years in the Phonological Awareness Literacy Screening (PALS) program?

Two 3-step hierarchical regression models were used to analyze Research Questions 5 and 6. In each hierarchical regression model, the personal characteristics of gender and age were entered in Step 1. Speed DIAL was entered at Step 2. In Step 3, the following control variables were entered: retention status (retained or not retained), special education status (did or did not receive special education services), and the number of years in the respective Title I math or reading program. The null hypotheses for Research Questions 5 and 6 were:

> $Ho5_1$ and $Ho6_1$: After controlling for the other independent variables in the model, there is no association between gender and Virginia Standards of Learning third-grade test scores.
>
> $Ho5_2$ and $Ho6_2$: After controlling for the other variables in the model, there is no association between age and Virginia Standards of Learning third-grade test scores.
>
> $Ho5_3$ and $Ho6_3$: After controlling for the other independent variables in the model, there is no association between Speed DIAL scores and Virginia Standards of Learning third-grade test scores.
>
> $Ho5_4$ and $Ho6_4$: After controlling for the other independent variables in the model, there is no association between retention status and Virginia Standards of Learning third-grade test scores.
>
> $Ho5_5$ and $Ho6_5$: After controlling for the other independent variables in the model, there is no association between special education status and Virginia Standards of Learning third-grade test scores.
>
> $Ho5_6$ and $Ho6_6$: After controlling for the other independent variables in the model, there is no association between the number of years in the Title I program and Virginia Standards of Learning third-grade test scores.

$Ho6_7$: After controlling for the other independent variables in the model, there is no association between the number of years in the PALS program and Virginia Standards of Learning third-grade total English test scores.

Summary

The study's results were derived from quantitative data obtained from Speed DIAL screening instrument scores and the Virginia Standards of Learning third-grade test scores. The data were obtained from students' permanent records. Descriptive and correlational statistics were used to analyze the data. Results from the analysis are presented in Chapter 4.

CHAPTER 4

RESULTS

The purpose of this study was to validate the psychometric integrity of the Speed DIAL by establishing its predictive validity. The Virginia Standards of Learning third-grade math score and the Virginia Standards of Learning third-grade total English score served as the outcome criteria for this research. The researcher also recognized that during the four- and one half-year span between the predictor variable and the outcome variable, many things might have happened in each student's academic life. Certainly, for this particular study it was impossible to go back four years and document the individual accommodations that were made daily in the classroom. However, it was possible to access archival data reporting when the students in the study were participants in intervention programs throughout their journey from preschool through third grade. The researcher included the intervening variables in the statistical analysis of the predictive validity of the Speed DIAL instrument. The intervening factors included in this study were: participation in Title I reading and math services, special education services, participation in phonological awareness program, and grade retention.

The data for this study came from three sources. The school system's Title I director provided the Speed DIAL scores, information about students' inclusion in Title I services, and grade retention. The curriculum and instruction department provided the Virginia Standards of Learning third-grade English and math scores as well as information about students' participation in the PALS Phonological Awareness program. The special education department identified the special education status for each student included in the study.

Although 288 students attended preschool during the years of 1999-2000 and 2000-2001, only 205 maintained continuous enrollment within the school system through third grade and participated in the Virginia Standards of Learning third grade end of year tests.

Descriptive Statistics

Tables 2 through 6 provide frequency and descriptive statistics for the study's population.

Table 2

Gender Status of the Population

Gender	*n*	%
Male	112	54.6
Female	93	45.4
Total	205	100

Table 3

Age Status of the Population

Age	*n*	%
Younger	98	47.8
Older	107	52.2
Total	205	100

Table 4

Retention Status of the Population

Retention Status	*n*	%
No	191	93.2
Yes	14	6.8
Total	205	100

Table 5

Special Education Status of the Population

Special Education Status	*n*	*%*
No	175	85.4
Yes	30	14.6
Total	205	100

Table 6

Descriptive Statistics for Variables With Interval Measurement

	N	*Mdn*	*M*	*SD*
Speed DIAL	205	94.0	93.18	10.104
Years in Title I Math	205	1.0	1.16	1.258
Years in Title I Reading	205	1.0	1.38	1.347
Years in PALS	205	.0	.30	.802
Total English Score	205	452.0	449.91	61.976
Math Score	202	509.5	511.93	69.409

Analysis of Data for Research Questions and Null Hypotheses

Research Question 1

Is there an association between Speed DIAL classification (*potential delay* versus *okay*) and pass-fail outcomes of the Virginia Standards of Learning third grade math test scores? The null hypothesis for this question was:

Ho1: There is no association between Speed DIAL classification and outcome on the Virginia Standards of Learning third grade math test scores.

It would be appropriate to use chi-square (X^2) to test the association between Speed DIAL classification and pass-fail outcomes on the Virginia Standards of Learning third-grade math test. However, to use chi-square with a crosstabulated 2 X 2 table, two assumptions must be met.

1. No more than 20% of the cells can have an expected frequency of less than five, and
2. The minimum expected frequency must be at least one.

There is a violation of the first assumption as 25% of the cells have an expected count less than five. Because of this violation of the assumption, the hypothesis was not tested using chi-square, but was tested using the descriptive statistics of the crosstabulated table as shown in Table 7.

Table 7

Cells for Calculating Overall Effectiveness and Indices of Sensitivity and Specificity

		Speed DIAL					
		Potential Delay		Okay		Total	
Math		*N*	%	*N*	%	*N*	%
	Fail	6	15.8	10	6.1	16	7.9
	Pass	32	82.4	154	93.9	186	92.1
	Total	38	100.0	164	100.0	202	100.0

Index of Sensitivity is the number of students identified as *potential delay* by Speed DIAL who later failed the Virginia Standards of Learning third-grade math test divided by the total number of students who failed the Virginia Standards of Learning third-grade math test. When using the Virginia Standards of Learning third-grade math test scores as the outcome criterion, the Sensitivity Index for Speed DIAL was 37.5%.

Index of Specificity is the number of children identified as *okay* by Speed DIAL who later passed the Virginia Standards of Learning third-grade math test divided by the number of students who passed the Virginia Standards of Learning third-grade math test. When using the Virginia Standards of Learning third-grade math test scores as the outcome criterion, the Specificity Index for Speed DIAL was 82.8%.

The overall effectiveness index is the number of children correctly identified as either *potential delay* or *okay* by Speed DIAL using the Virginia Standards of Learning third-grade math test score as the outcome criteria divided by the number of children screened using Speed DIAL. Again, using the Virginia Standards of Learning third-grade math test score as the outcome criterion, the overall effectiveness index for Speed DIAL was 79%.

Two other indices reflect the percentage of correct identifications of *potential delay* and *okay* using Speed DIAL. For students originally identified as *potential delay* using Speed DIAL, the percentage of correct identifications using the Virginia Standards of Learning third-grade math test scores as the outcome criterion was 15.7%. For students originally identified as *okay* using Speed DIAL, the percentage of correct identifications using the Virginia Standards of Learning third-grade math test scores as the outcome criterion was 94%.

Research Question 2

Is there an association between Speed DIAL classification (*potential delay* versus *okay*) and pass-fail outcomes of the Virginia Standards of Learning third-grade total English test scores? The null hypothesis for this question was:

Ho2: There is no association between Speed DIAL classification and outcome on the Virginia Standards of Learning third grade total English test scores.

Chi-square (X^2) was used to test the association between Speed DIAL classification and pass-fail outcomes on the Virginia Standards of Learning third-grade total English test. The analysis showed there was a significant association between Speed DIAL classification and pass-

fail outcome on the Virginia Standards of Learning third-grade total English test, X^2 (1) = 8.233, p = .004. Therefore, the null hypothesis was rejected.

Table 8

Cells for Calculating Overall Effectiveness and Indices of Sensitivity and Specificity

		Speed DIAL					
		Potential Delay		Okay		Total	
Total English		*N*	%	*N*	%	*N*	%
	Fail	14	35.0	25	15.2	39	19.0
	Pass	26	65.0	140	84.8	166	81.0
	Total	40	100.0	165	100.0	205	100.0

As shown in Table 8, Index of Sensitivity is the number of students identified as *potential delay* by Speed DIAL who later failed the Virginia Standards of Learning third-grade total English test divided by the number of students who failed the Virginia Standards of Learning third-grade total English test. When using the Virginia Standards of Learning third-grade total English test scores as the outcome criterion, the Sensitivity Index for Speed DIAL was 35.9%.

Index of Specificity is the number of children identified as *okay* by Speed DIAL who later passed the Virginia Standards of Learning third-grade total English test divided by the number of students who passed the Virginia Standards of Learning third-grade total English test. When using the Virginia Standards of Learning third-grade total English test scores as the outcome criterion, the Specificity Index for Speed DIAL was 84.9%.

The overall effectiveness index is the number of children correctly identified as either *potential delay* or *okay* by Speed DIAL using the Virginia Standards of Learning third-grade

total English test score as the outcome criteria divided by the number of children screened using Speed DIAL. Again, using the Virginia Standards of Learning third-grade total English test score as the outcome criterion, the overall effectiveness index for Speed DIAL was 75%.

Two other indices reflect the percentage of correct identifications of *potential delay* and *okay* using Speed DIAL. For students originally identified as *potential delay* using Speed DIAL, the percentage of correct identifications using the Virginia Standards of Learning third-grade total English test scores as the outcome criterion was 35%. For students originally identified as *okay* using Speed DIAL, the percentage of correct identifications using the Virginia Standards of Learning third-grade total English test scores as the outcome criterion was 85%.

Research Question 3

Is there an association between gender, age, and Speed DIAL scores? The null hypotheses for this question were:

$Ho3_1$: There is no association between gender and Speed DIAL scores.

$Ho3_2$: There is no association between age and Speed DIAL scores.

$Ho3_3$: There is no significant gender by age interaction.

A two-way analysis of variance was conducted to evaluate the associations between gender and age and Speed DIAL scores. The dependent variable was Speed DIAL scores. The independent variables were gender and age (younger and older students). In addition, a two-way interaction term for gender by age was tested.

The gender by age interaction was not significant, F (1,201) = .882, p = .35, partial η^2 = .004. Therefore, the null hypothesis for gender by age interaction was retained and it was appropriate to proceed with the examination of the main effects of gender and age.

There was an association between gender and Speed DIAL scores, F (1,201) = 7.381, p = .007. Therefore, the null hypothesis for gender was rejected. The Speed DIAL mean for females (M = 95.3 SD = 8.94) was almost 4 points higher than the mean for males (M = 91.4 SD = 10.7).

The strength of the association, as measured by the partial η^2, was small (.035). Finally, there was no association between age and Speed DIAL scores, F (1,201) = 1.562, p = .213. The partial η^2 for age was small (.008). Therefore, the null hypothesis for age was retained. The means and standard deviations for gender and age are reported in Table 9.

Table 9

Means and Standard Deviations for Speed DIAL Scores by Gender and Age

Gender	Age	*n*	*M*	*SD*
Male	Younger	50	91.66	10.22
	Older	62	91.23	11.15
	Total	112	91.42	10.70
Female	Younger	48	96.77	8.50
	Older	45	93.71	9.23
	Total	93	95.29	8.94
Total	Younger	98	94.16	9.71
	Older	107	92.27	10.42
	Total	205	93.18	10.10

Research Question 4

Is there an association between Speed DIAL scores and the intervening variables: retention status, special education status, and number of years of participation in Title I math, Title I reading, and PALS?

Pearson's correlations were used to analyze this research question. The null hypotheses were:

$Ho4_1$: There is no association between Speed DIAL scores and retention status.

$Ho4_2$: There is no association between Speed DIAL scores and special education status.

$Ho4_3$: There is no association between Speed DIAL scores and the number of years of participation in Title I math.

$Ho4_4$: There is no association between Speed DIAL scores and the number of years of participation in Title I reading.

$Ho4_5$: There is no association between Speed DIAL scores and the number of years of participation in the PALS reading program.

Table 10

Correlations Between Speed DIAL and Intervention Variables

Intervention	*N*	Speed DIAL *r*	Speed DIAL r^2
Retention Status	205	-.25**	.063
Special Education Status	205	-.17*	.029
Years in Title I Math	167	-.33**	.109
Years in Title I Reading	167	-.32**	.102
Years in PALS	151	-.27**	.073

**Correlation is significant at the 0.01 level
*Correlation is significant at the 0.05 level

As shown in Table 10, Pearson's correlation was used to analyze the association between Speed DIAL scores and the documented interventions. Those interventions were retention

status, special education status, number of years in Title I math program, number of years in Title I reading program, and number of years in PALS program. Pearson's correlation can range from -1 to +1. Correlations of zero represent no association and associations are considered stronger as they get closer to -1 or +1. As with most research studies, there are many variables that cannot be controlled for, such as home environment. What is considered a small or large correlation is dependent upon the particulars special to each area of investigation. Because of this, any correlation above or below zero can be considered as indicative of an association. In this case, the correlations while not considered strong, do indicate negative associations between Speed DIAL and the intervention variables. The analysis indicated that lower Speed DIAL scores resulted in longer years of participation in Title I math services. In fact, 11% (r^2 = .109) of the variance in years participating in Title I math was accounted for by Speed DIAL scores. Lower Speed DIAL scores also indicated longer years of participation in Title I reading and PALS program as well as a greater chance of being retained and/or served by Special Education services. Speed DIAL accounted for 6.4% of the variance in retention status (r^2 = .063), 2.9% of the variance in Special Education status (r^2 = .029), 7.3% of the variance in years participating in the PALS program (r^2 = .073) and 10.2% of the variance in number of years in Title I reading (r^2 = .102). While these percentages of variance accounted for certainly would not be considered great, they are both statistically and substantively significant. All five null hypotheses were rejected for this question.

Research Question 5

Is there an association between the Virginia Standards of Learning third-grade math test scores and gender, age, Speed DIAL scores, retention status, special education status, and number of years in the Title I math program?

The null hypotheses for this question were:

$Ho5_1$: After controlling for the other independent variables in the model, there is no association between gender and Virginia Standards of Learning third-grade math scores.

$Ho5_2$: After controlling for the other variables in the model, there is no association between age and Virginia Standards of Learning third-grade math scores.

$Ho5_3$: After controlling for the other independent variables in the model, there is no association between Speed DIAL scores and Virginia Standards of Learning third-grade math scores.

$Ho5_4$: After controlling for the other independent variables in the model, there is no association between retention status and Virginia Standards of Learning third-grade math scores.

$Ho5_5$: After controlling for the other independent variables in the model, there is no association between special education status and Virginia Standards of Learning third-grade math scores.

$Ho5_6$: After controlling for the other independent variables in the model, there is no association between the number of years in the Title I math program and Virginia Standards of Learning third-grade math scores.

A hierarchical multiple regression was used to analyze the associations between the independent variables and the dependent variable. For this regression model, the analysis was limited to students who attended schools with the Title I math program ($N = 164$).

The dependent variable was the Virginia Standards of Learning third-grade math test score. The independent dichotomous variables were: gender, age (younger versus older students), retention status (not retained versus retained) and special education status (did not receive special education services versus did receive special education services). These dichotomous variables were dummy coded 0 and 1. Speed DIAL scores and the number of years in the Title I math program were the two continuous independent variables in the model.

A preliminary analysis of the residuals was conducted to determine the aptness of the regression model. A one-sample Kolmogorov-Smirnov test was used to determine whether or not the residuals deviated from a normal distribution. The one-sample Kolmogorov-Smirnov test was not significant (p = .69). Therefore, the assumption of a normal distribution of the residuals was met. Figure 1 shows the distribution of the residuals.

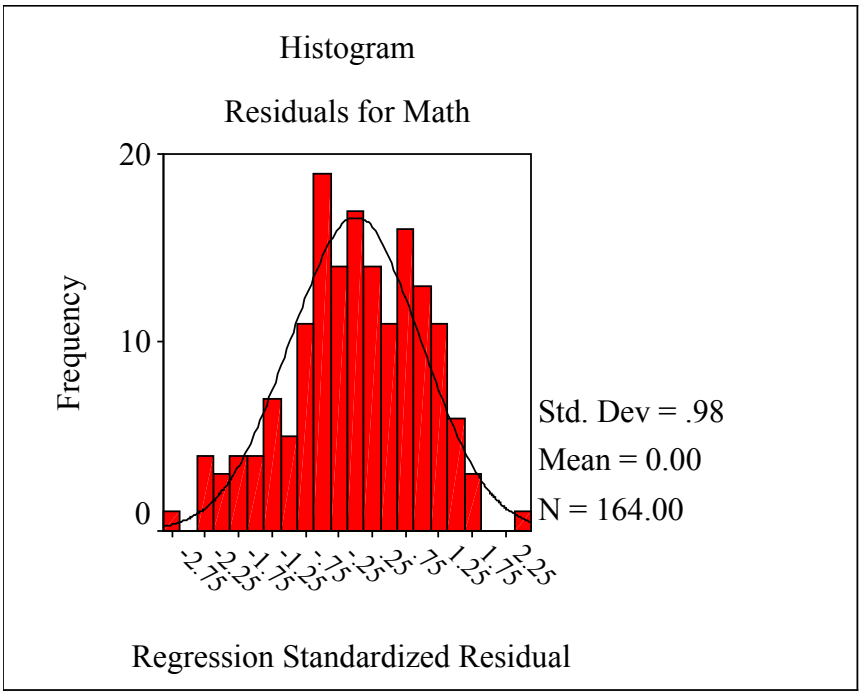

Figure 1. Histogram of the Standardized Residuals of the Regression Model for Math

A visual inspection of the normal probability plot of the standardized residuals was used also to check for normality. If the variable is normally distributed, the plotted points form a straight diagonal line. As shown in Figure 2, there did not appear to be a deviation from normality.

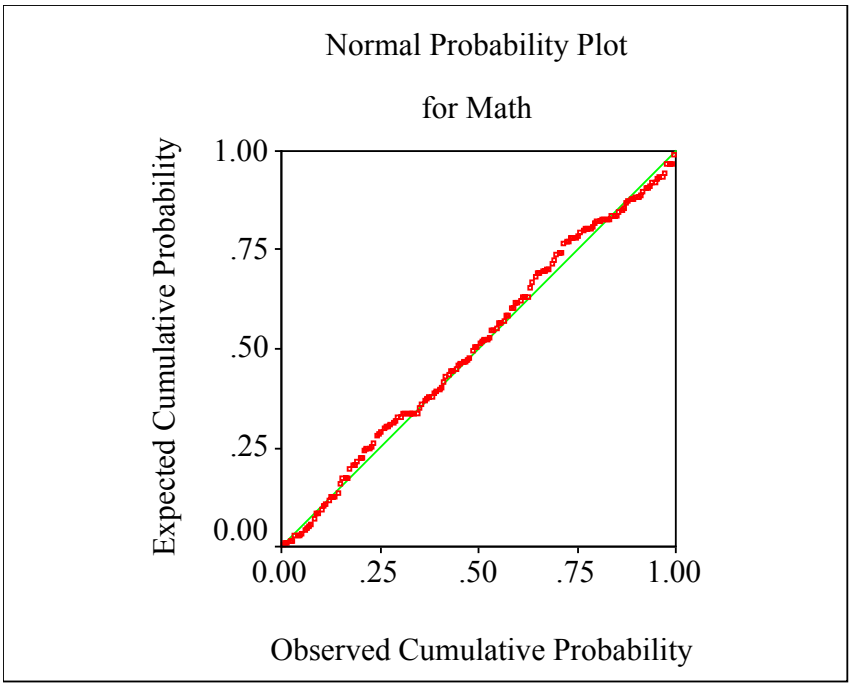

Figure 2. Normal Probability Plot of the Standardized Residuals of the Regression Model for Math

Finally, the scatterplot of the standardized residuals regressed on the standardized predicted values is shown in Figure 3. If the regression model is appropriate, the scatterplot should show no discernible pattern. Visual inspection of this plot raised suspicion that there may be a negative association between the standardized residuals and the standardized predicted values. Therefore, further investigation was warranted.

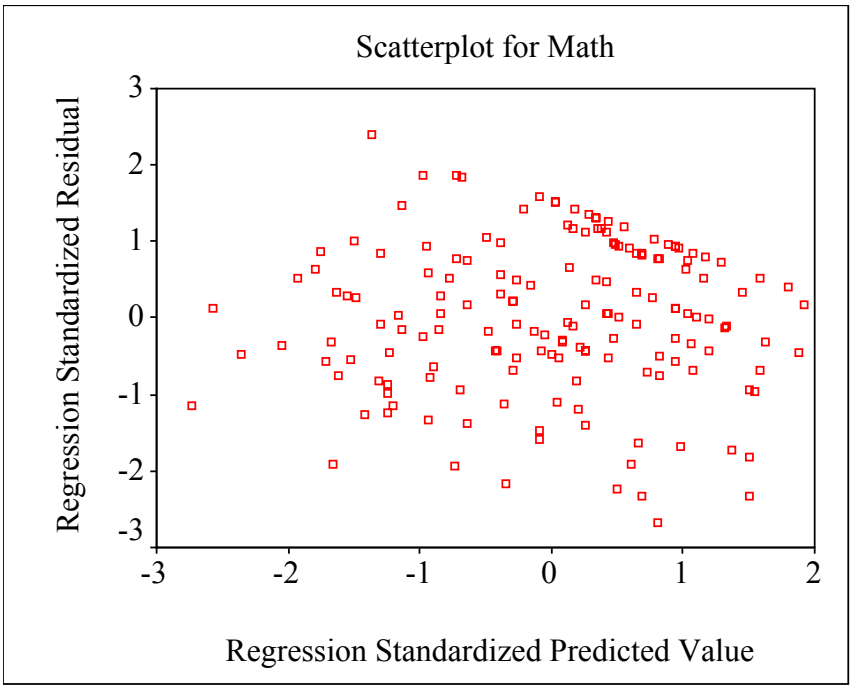

Figure 3. Scatterplot of the Standardized Residuals and Standardized Predicted Values for the Regression Model for Math

To determine whether or not there was a linear and/or curvilinear association between the standardized residuals and standardized predicted values, I used the SPSS regression curve estimation procedure. For this regression model the dependent variable was the standardized residuals and the two independent variables were: the standardized predicted values used to test for linearity (the linear term) and the squared standardized predicted values used to test for curvilinearity (the quadratic term). The results showed that neither the linear (p = .56) nor the quadratic term (p = .06) was significant. However, while the quadratic term was not significant, the plotted line was slightly curved as shown in Figure 4. Because the quadratic term accounted for only 2% of the variance in the standardized residuals (r^2 change = .021), it was reasonable to conclude the regression model was appropriate.

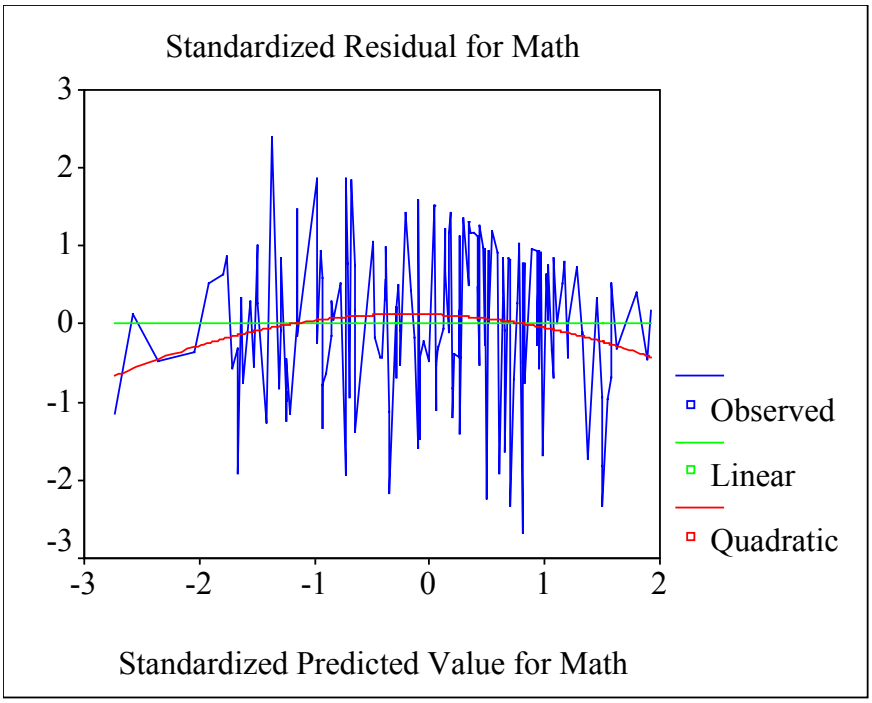

Figure 4. Scatterplot of Standardized Residuals and Standardized Predicted Values

Correlations among the independent variables and the dependent variable are shown in Table 11.

Table 11

Correlations Among the Predictors and Math Scores

	Math	Gender	Age	Speed DIAL	Retain	Spec. Educ.	Title Math I
Math Score	1.000						
Gender	-.119	1.000					
Age	-.001	-.053	1.000				
Speed DIAL	.334*	.201*	-.076	1.000			
Retained	-.199*	-.053	-.087	-.288*	1.000		
Spec. Educ.	-.219*	-.207*	-.008	-.177*	.224*	1.000	
Title I Math	-.446*	-.064	-.045	-.340*	.299*	.201*	1.000

* Significant at the .01 level.

The results of the hierarchical regression model are shown in Table 12.

Table 12

Hierarchical Multiple Regression of the Effects of Personal Characteristics, Speed DIAL Scores, and Intervention Variables on Math Scores

Independent Variables	B	β	Part r^2	*p*
Step 1				
Gender	-17.18	-.12	.014	.13
Age	-1.08	-.01	<.001	.92

Table 12 (continued)

Independent Variables	B	β	Part r^2	*p*
Step 2				
Gender	-27.80	-.19	.036	.01**
Age	2.38	.02	<.001	.82
Speed DIAL	2.66	.37	.134	<.01**
Step 3				
Gender	-31.60	-.22	.045	<.01**
Age	-1.88	-.01	<.001	.85
Speed DIAL	1.63	.23	.042	<.01**
Retention	-1.88	-.01	<.001	.92
Spec. Educ.	-31.12	-.15	.021	.03*
Title I Math	-19.85	-.350	.101	<.01**

Step 1: F (2, 161) = 1.17, R^2 = .014, p = .313
Step 2: F (3, 160) = 9.28, R^2 = .148, p < .001*
Step 3: F (6, 157) = 10.89, R^2 = .294, p < .001*
*Significant at the .05 level
**Significant at the .01 level

In Step 1, gender and age were entered in the model. The overall regression model at this step was not significant, F (2, 161) = 1.17, R^2 = .014, p = .31. Gender and age together accounted for only 1.4% of the variance in math scores. The regression coefficients for gender (β = - .12, p = .13) and age (β = -.01, p = .92) were not significant.

Speed DIAL scores were entered in the regression model at Step 2. The model containing gender, age and Speed DIAL scores was significant, F (3, 160) = 9.28,

R^2 = .148, p < .01, with the three variables together accounting for 14.8% of the variance in math. The regression coefficient for age remained nonsignificant at Step 2 (β = .02, p = .82). However, the coefficient for gender was significant (β = -.19, p = .01) with males having higher scores on math than females. The regression coefficient for Speed DIAL scores was also significant (β = .374, p < .01). The part correlation square (part^2) for Speed DIAL, which is equivalent to the r^2 change when the variable is entered in the model last, showed that of the 14.8% of the total variance in math accounted for by gender, age, and Speed DIAL scores together, 13.4% was accounted for by Speed DIAL scores. Clearly, the majority of explained variance at Step 2 was attributed to the inclusion of Speed DIAL scores in the model.

Step 3 of the hierarchical model included age, gender, Speed DIAL scores, and other control variables: retention status, special education status, and the number of years in the Title I math program. The overall model at the final Step 3 was significant, F (6, 157) = 10.89, R^2 = .294, p < .01. All six independent variables together accounted for 29.4% of the variance in math scores. The R^2 change between steps two and three (R^2 change = .146) showed that the inclusion of retention status, special education status, and the number of years in Title I math in the model contributed an additional 14.6% to the explained variance in math scores, variance not accounted for by age, gender, and Speed DIAL scores in Step 2.

At Step 3, the regression coefficients for age (β = -.01, p = .85) and retention status (β = -.01, p = .92) were not significant. The coefficient for gender was significant (β = -.22, p = <.01) showing males had higher scores on math than did females. In addition, the regression coefficient for special education status was significant, (β = -.15, p = .03). Students who had not received special education services had higher scores on math than did students who received these services. The number of years in Title I math was also statistically significant (β = -.35, p = <.01) as was Speed DIAL scores ((β = .23, p = <.01).

In the third and final step of the hierarchical regression model, the examination of the part correlation squared coefficients (part r^2) showed that the number of years in Title I math

contributed the most to the total variance explained by the model (part r^2 = .101). In other words, when entered last in the model, the number of years in Title I math uniquely contributed an additional 10.1% to the variance explained. However, the association between the number of years in Title I math and math scores was negative. The unique contribution to the total variance explained by gender (part r^2 = .045) was 4.5% while Speed DIAL scores (part r^2 = .042) uniquely contributed 4.2% to the variance explained.

The summary of the findings for the null hypotheses include: The null hypotheses for age and retention status were retained. The null hypothesis for gender was rejected with males having higher math scores than females. The null hypothesis for special education status was also rejected with students who had not received special education services having higher math scores than students who received the services. The null hypothesis for Speed DIAL scores was rejected: Students with higher scores on Speed DIAL had higher scores on math. Finally, the null hypothesis for the number of years in Title I math was rejected showing that the longer students had participated in the Title I math program, the lower were their math scores.

With regard to the predictive value of Speed DIAL scores on the Virginia Standards of Learning math scores, it was found that in Step 1 Speed DIAL scores accounted for 13.4% of the 14.8% total variance accounted for by age, gender, and Speed DIAL scores. When the other control variables (retention status, special education status, and the number of years in the Title I math program) were included in the model at Step 3, the unique contribution of Speed DIAL scores to the total variance explained (29.4%) was 4.2%. However, of greater importance was the change in the *Beta* coefficients for Speed DIAL between Steps 2 and 3 in the regression model. The *Beta* coefficient for Speed DIAL in Step 2 was .37. In Step 3 the coefficient for Speed DIAL after controlling for age, gender and the three intervention variables (Special Education Status, Retention Status and number of years of participation in Title I Math) dropped to .23. The substantive importance of the reduction in the magnitude of the *Beta* coefficients for Speed DIAL between Step 2 to Step 3 will be addressed in Chapter 5.

Research Question 6

Is there an association between the Virginia Standards of Learning third-grade total English test scores and gender, age, Speed DIAL scores, retention status, special education status, number of years in the Title I reading program, and number of years in the Phonological Awareness Literacy Screening (PALS) program?

The null hypotheses for this question were:

$Ho6_1$: After controlling for the other independent variables in the model, there is no association between gender and Virginia Standards of Learning third-grade total English scores.

$Ho6_2$: After controlling for the other variables in the model, there is no association between age and Virginia Standards of Learning third-grade total English scores.

$Ho6_3$: After controlling for the other independent variables in the model, there is no association between Speed DIAL scores and Virginia Standards of Learning third-grade total English scores.

$Ho6_4$: After controlling for the other independent variables in the model, there is no association between retention status and Virginia Standards of Learning third-grade total English scores.

$Ho6_5$: After controlling for the other independent variables in the model, there is no association between special education Status and Virginia Standards of Learning third-grade total English scores.

$Ho6_6$: After controlling for the other independent variables in the model, there is no association between the number of years in the Title I reading program and Virginia Standards of Learning third-grade total English scores.

$Ho6_7$: After controlling for the other independent variables in the model, there is no association between the number of years in the PALS phonological awareness program and Virginia Standards of Learning third-grade total English scores.

A hierarchical multiple regression was used to analyze the associations between the independent variables and the dependent variable. For this regression model, the analysis was limited to students who attended schools with the PALS program and the Title I reading program ($N = 151$).

The dependent variable was the Virginia Standards of Learning third-grade total English test score. The independent dichotomous variables were: gender, age (younger versus older students), retention status (not retained versus retained), and special education status (did not receive special education services versus did receive special education services). These dichotomous variables were dummy coded 0 and 1. Speed DIAL scores, number of years in the Title I reading, and number of years in the PALS program were the three continuous independent variables in the model.

A preliminary analysis of the residuals was conducted to determine the aptness of the regression model. A one-sample Kolmogorov-Smirnov test was used to determine whether or not the residuals deviated from a normal distribution. The one-sample Kolmogorov-Smirnov test was not significant ($p = .40$). Therefore, the assumption of a normal distribution of the residuals was met. Figure 5 shows the distribution of the residuals.

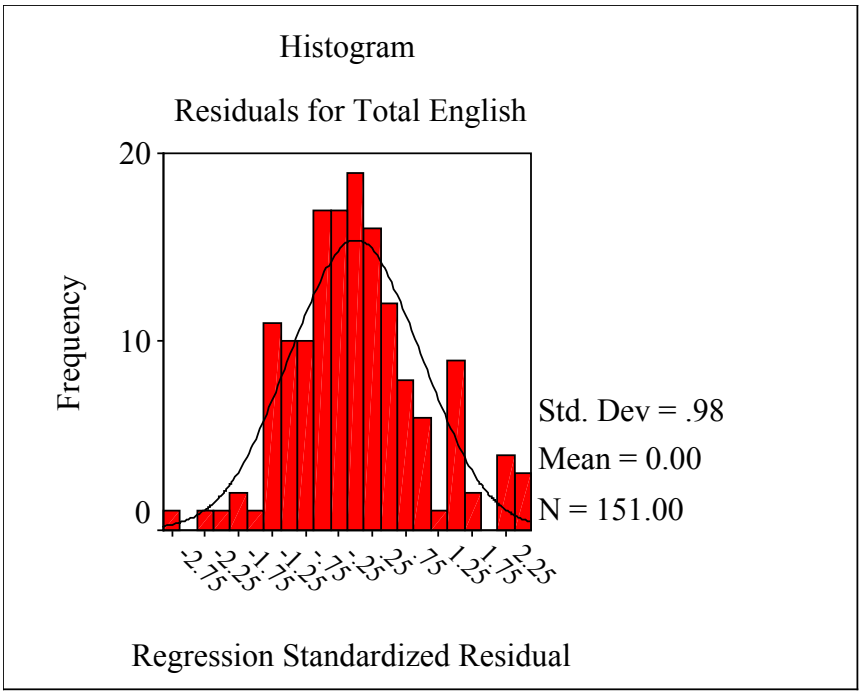

Figure 5. Histogram of the Standardized Residuals for Total English Scores

A visual inspection of the normal probability plot of the standardized residuals was used also to check for normality. If the variable is normally distributed, the plotted points form a straight diagonal line. As shown in Figure 6, there did not appear to be a deviation from normality.

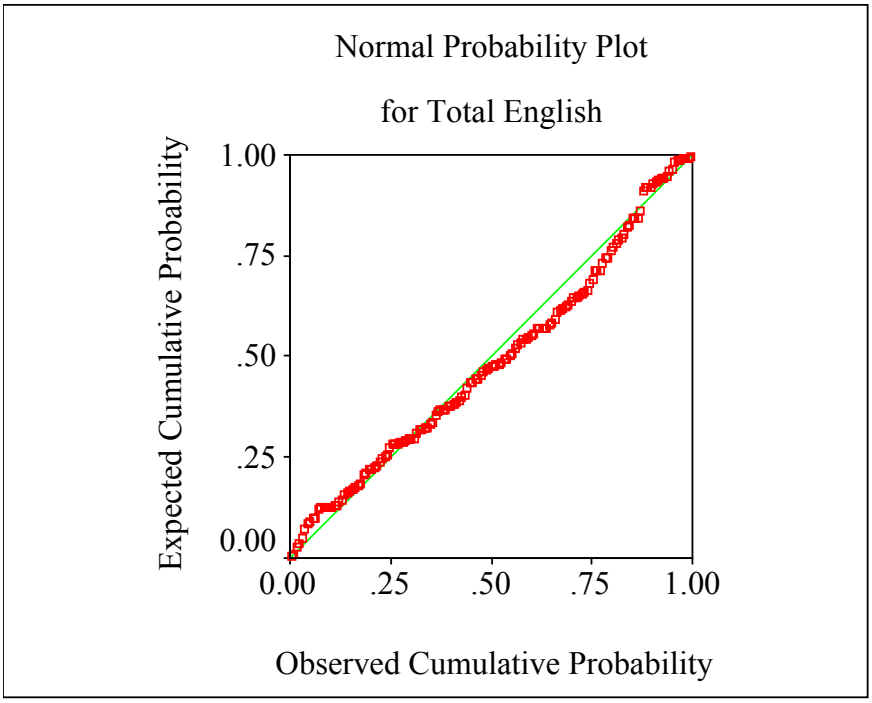

Figure 6. Normal Probability Plot of the Standardized Residuals of the Regression Model for Total English

Finally, the scatterplot of the standardized residuals regressed on the standardized predicted values is shown in Figure 7. If the regression model is appropriate, the scatterplot should show no discernible pattern. Visual inspection of this plot raised suspicion that there may be a linear or curvilinear association between the standardized residuals and the standardized predicted values. As a precaution, further investigation was warranted.

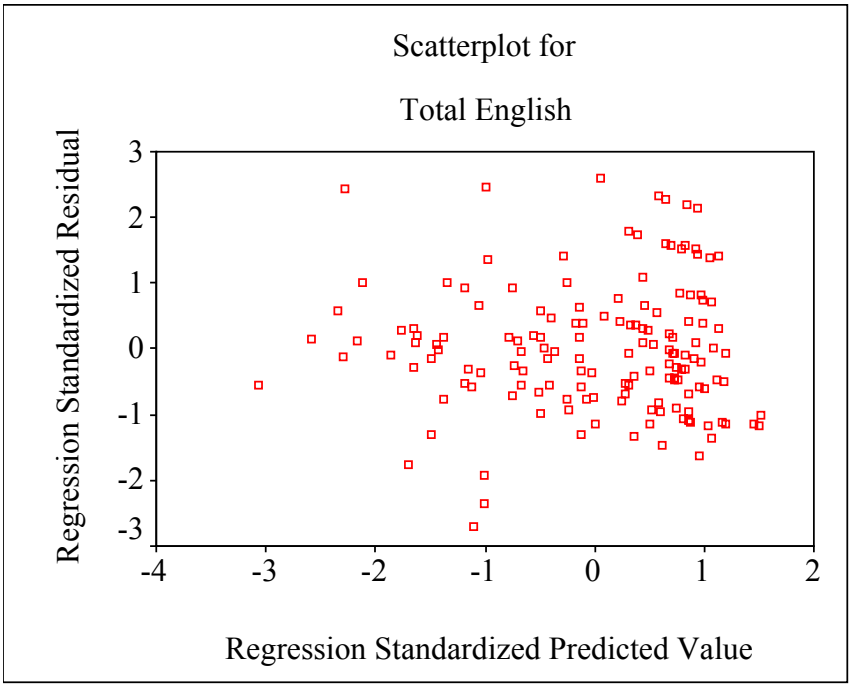

Figure 7. Scatterplot of the Standardized Residuals and Standardized Predicted Values for the Regression Model for Total English

To determine whether or not there was a linear and/or curvilinear association between the standardized residuals and standardized predicted values, I used the SPSS regression curve estimation procedure. For this regression model the dependent variable was the standardized residuals and the two independent variables were the standardized predicted values used to test for linearity (the linear term) and the squared standardized predicted values used to test for curvilinearity (the quadratic term). The results showed that neither the linear (p = .99) nor the quadratic term (p = .99) was significant and, therefore, indicated there were no violations of the assumptions of regression.

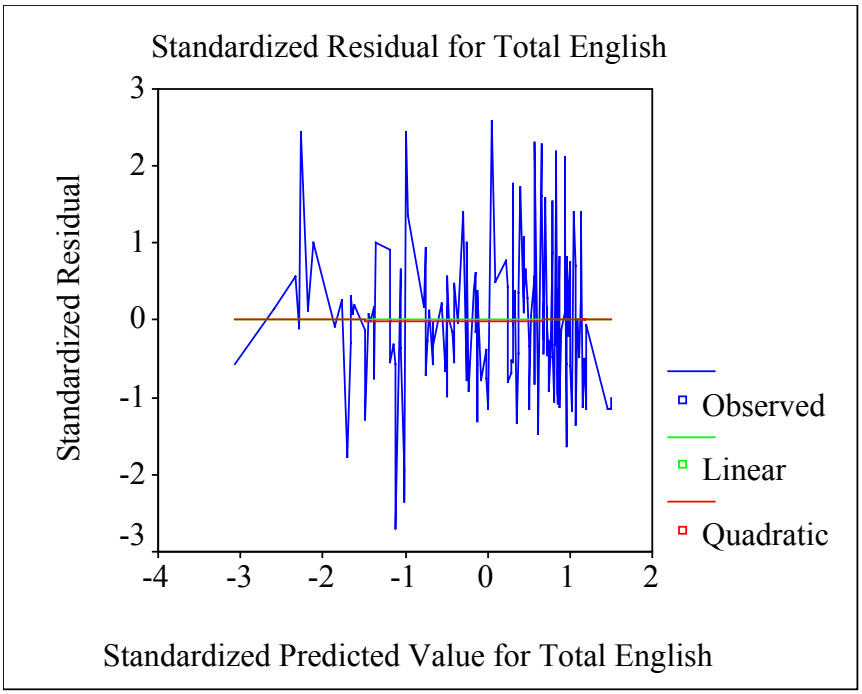

Figure 8. Plot of Linear and Quadratic Functions of the Standardized Residuals and Standardized Predicted Values for Total English

Correlations among the independent variables and the dependent variable are shown in Table 13

Table 13

Correlations Among the Predictors and Total English Scores

	Total English Score	Gender	Age	Speed DIAL	Retain	Spec. Educ.	Title I Reading	PALS
Gender	.11	1.00						
Age	-.07	-.07	1.00					
Speed DIAL	.31**	.19**	-.10	1.00				
Retained	-.21**	-.04	-.11	-.28**	1.00			
Spec. Educ.	-.31**	-.25**	.06	-.27**	.21**	1.00		
Title I Read	-.31**	-.05	-.01	-.30**	.21**	.08	1.00	
PALS	-.28**	.01	-.03	-.27**	.27**	.10	.43**	1.00

*Significant at the .05 level
**Significant at the .01 level

The results of the hierarchical regression model are shown in Table 14

Table 14

Hierarchical Multiple Regression of the Effects of Personal Characteristics, Speed DIAL Scores, and Intervention Variables on Total English Scores

Independent Variables	B	β	Part r^2	*p*
Step 1				
Gender	13.37	.10	.010	.22
Age	-8.12	-.06	.004	.45
Step 2				
Gender	6.19	.05	.002	.56
Age	-4.54	-.04	.001	.66
Speed DIAL	1.97	.30	.087	<.01**
Step 3				
Gender	1.33	.01	<.001	.90
Age	-6.90	-.05	.003	.48
Speed DIAL	.94	.14	.016	.09
Retention	-13.65	-.06	.003	.48
Spec. Educ.	-39.53	-.22	.043	<.01**
Title I Reading	-8.86	-.18	.025	<.04*
PALS Program	-9.03	-.13	.012	.14

Step 1: F (2, 148) = 1.124, R^2 = .015, p = .328
Step 2: F (3, 147) = 5.547, R^2 = .102, p = .001**
Step 3: F (7, 143) = 5.720, R^2 = .219, p < .001**
*Significant at the .05 level
**Significant at the .01 level

In Step 1, gender and age were entered in the model. The overall regression model at this step was not significant, F (2, 148) = 1.124, R^2 = .015, p = .328. Gender and age together accounted for only 1.5% of the variance in total English scores. The regression coefficients for gender (β = - .10, p = .22) and age (β = -.06, p = .45) were not significant.

Speed DIAL scores were entered in the regression model at Step 2. The model containing gender, age, and Speed DIAL scores was significant, F (3, 147) = 5.547, R^2 = .102, p = .001, with the three variables together accounting for 10.2% of the variance in total English. The regression coefficient for age remained nonsignificant at Step 2 (β = -.04, p = .66). The coefficient for gender was nonsignificant (β = .05, p = .56). The regression coefficient for Speed DIAL scores was significant (β = .301, p < .01). The part correlation square for Speed DIAL, which is equivalent to the r^2 change when the variable is entered in the model last, showed that of the 10.2% of the total variance in total English accounted for by gender, age and Speed DIAL scores together, 8.7% was accounted for by Speed DIAL scores. Clearly, the majority of explained variance at Step 2 was attributed to the inclusion of Speed DIAL scores in the model.

Step 3 of the hierarchical model included age, gender, Speed DIAL scores, and other control variables: retention status, special education status, the number of years in the Title I reading program, and the number of years in the PALS program. The overall model at the final Step 3 was significant, F (7.143) = 5.720, R^2 = .219, p < .001. All seven independent variables together accounted for 21.9% of the variance in total English scores. The R^2 change between steps two and three (R^2 change = .117) showed that the inclusion of retention status, special education status, the number of years in Title I reading, and the number of years in the PALS program in the model contributed an additional 11.7% to the explained variance in total English scores, variance not accounted for by age, gender and Speed DIAL scores in Step 2.

At step 3, the regression coefficients for age (β = -.05, p = .48) and retention status (β = -.06, p = .48) were not significant. The coefficient for gender was not significant (β = .01, p =

.90). The regression coefficient for special education status was significant, (β = .22, p = <.01). Students who had not received special education services had higher scores on total English than students who had received these services. The number of years in Title I reading was also statistically significant (β = -.178, p = .04). Students who had not received Title I reading services had higher scores on the total English than did students who received these services. The number of years in the PALS program was nonsignificant (β = -.126, p = .14) as was Speed DIAL scores (β = .143, p = .09). In step 3, Speed DIAL was no longer statistically significant (p=.09). The zero-order correlation or original association between Speed DIAL and total English scores was .31. In step 3, the partial correlation between Speed DIAL and total English, after all the other variables were controlled, dropped to .143 so the original association between the two diminished.

In the third and final step of the hierarchical regression model, the examination of the partial correlation squared coefficients (part r^2) showed that special education status contributed the most to the variance explained by the model (part r^2 = .043). In other words, when entered in the model last, special education status uniquely contributed an additional 4.3% to the variance explained. However, the association between special education status and total English scores was negative. The unique contribution to the total variance explained by number of years in Title I reading (part r^2 = .025) was 2.5% whereas Speed DIAL scores (part r^2 = .016) uniquely contributed only 1.6% to the variance explained.

The summary of the findings for the null hypotheses are: The null hypotheses for gender, age, Speed DIAL scores, retention status, and years in PALS program were retained. The null hypothesis for special education status was rejected with special education students having lower total English scores than those students not receiving special education services. Finally, the null hypothesis for the number of years in Title I reading was rejected showing that the longer students had participated in the Title I reading program, the lower was the total English score.

With regard to the predictive value of Speed DIAL scores on the Virginia Standards of Learning total English scores, it was found that in Step 1 Speed DIAL scores accounted for 8.7% of the 10.2% variance accounted for by age, gender, and Speed DIAL scores. When the other control variables (retention status, special education status, the number of years in the Title I reading program, and number of years in PALS program) were included in the model at Step 3, the unique contribution of Speed DIAL scores to the total variance explained (21.9%) was 1.6%, which was not statistically significant. Of greater importance was the change in the regression coefficient for Speed DIAL between Steps 2 and 3. At Step 2, the *Beta* coefficient for Speed DIAL was .30, which was significant (p <.01). However, at Step 3, the coefficient dropped to .14 and was not statistically significant (p = .09). The implications of these findings will be addressed in Chapter 5.

Summary

The purpose of this study was to validate the psychometric integrity of the Speed DIAL by establishing its predictive validity. To do this, research was conducted to determine what, if any, was the association between Speed DIAL scores and scores on the Virginia Standards of Learning third-grade tests. Six research questions were developed for the study. Several statistical tests were used to analyze the data. They included chi-square, Pearson's correlation, an analysis of variance (ANOVA), and a 3-step hierarchical regression. A crosstabulated matrix was also used with descriptive statistics to measure several indices related to Speed DIAL's predictive properties and association to the outcome criteria. Intervening variables occurring after the predictor variable but before the outcome variable were also included in the analysis, as were the factors of age and gender.

The findings in chapter four suggest that Speed DIAL does possess predictive qualities. There was a moderate correlation between Speed DIAL scores and Virginia Standards of Learning third-grade test scores. Speed DIAL's overall effectiveness rating exceeded 75%.

Females scored higher than did males on the Speed DIAL, and there was a negative association between Speed DIAL and the documented intervening variables.

CHAPTER 5

SUMMARY OF FINDINGS, CONCLUSIONS, AND RECOMMENDATIONS

The primary purpose of this study was to establish and evaluate the predictive validity of the Speed DIAL screening instrument. A screening instrument possesses predictive validity if it is able to distinguish between those students with potential learning disabilities and those students who do not have learning difficulties. Predictive validity of an instrument is established by correlating the results of the predictor variable with the results of an outcome variable (Bredekamp & Shepard, 1989). For this study, the Speed DIAL classifications of okay and potential delay were compared with the pass-fail designations of the outcome criteria, the Virginia Standards of Learning third-grade math and total English tests. In addition, actual scores on both the Speed DIAL and Virginia Standards of Learning third-grade math and total English tests were correlated using multiple regression models to determine if there was an association between the two variables. A secondary purpose of this study was to determine the association between age, gender, and the two assessment variables. Finally, data were collected to interpret the associations of several intervening variables relative to the independent and dependent variables.

Summary of Findings

The analysis focused on six research questions. The independent variables included in this study were Speed DIAL scores, gender, age, and several intervening variables including retention status, special education status, and number of years participating in Title I math, Title I reading, and the PALS Phonological Awareness Program. The dependent variables were the scores reported for third-grade students on the Virginia Standards of Learning third-grade math and total English tests. The population consisted of 205 students. These 205 students were

assessed using the Speed DIAL screening instrument while enrolled in a preschool program during the two school years of 1999-2000 and 2000-2001 and who maintained enrollment within the school district and were administered the Virginia Standards of Learning third-grade math and total English tests.

Research Question 1

Is there an association between Speed DIAL classification (potential delay vs. okay) and pass-fail outcomes of the Virginia Standards of Learning third-grade math test scores?

It was not possible to test the association between Speed DIAL classification and pass-fail outcomes on the Virginia Standards of Learning third-grade math test using chi-square because an assumption could not be met. The unmet assumption was that no more than 20% of the cells in the cross tabulated table could have an expected frequency of less than 5. Therefore, the hypothesis was not tested.

Research Question 2

Is there an association between Speed DIAL classification (potential delay vs. okay) and pass-fail outcomes of the Virginia Standards of Learning third grade total English test scores?

The null hypothesis for question two was rejected as the chi-square analysis showed there was a significant association between Speed DIAL classification and pass-fail outcomes on the Virginia Standards of Learning total English tests, x^2 (1)=8.233, p=.004.

Findings Related to Cross Tabulated Matrix Used to Establish Indices of Sensitivity, Specificity, and Overall Effectiveness

In the best of all circumstances, a school using Speed DIAL would accurately classify students as either okay or potential delay. However, complete accuracy is impossible because of factors that screeners have no control over. Czudnowski and Goldenberg (1998) explained:

> No matter how careful examiners are or how well constructed the test, there will be some error involved in the process. Some children identified in screening as okay may actually be potential delay while others identified as potential delay may actually be okay. (p.50)

Indices were calculated using descriptive statistics to determine the effectiveness of the Speed DIAL screening assessment.

When using the Virginia Standards of Learning third-grade math test scores as the outcome criterion, Speed DIAL's overall effectiveness index was 79%. One hundred sixty students out of 202 were correctly classified as either okay or potential delay by Speed DIAL when the Virginia Standards of Learning third-grade math test scores were used as the follow-up, or outcome criterion. When using the Virginia Standards of Learning third-grade total English scores as the outcome criterion, Speed DIAL's overall effectiveness index was 75%. Out of 205, 154 students were correctly classified as either okay or potential delay by Speed DIAL when the Virginia Standards of Learning third-grade total English tests scores were used as the follow-up, or outcome criterion.

Indices of sensitivity and specificity can be used to analyze the number of under-referrals and over-referrals. The index of sensitivity is the percentage of students who failed the Virginia Standards of Learning third-grade tests and were correctly classified using Speed DIAL as potential delay four years earlier. When using the Virginia Standards of Learning third-grade math test score as the outcome criterion, Speed DIAL's index of sensitivity was 37.5%. Sixteen students were classified as failing the Virginia Standards of Learning third-grade math test. Among these 16 students, 6 were originally classified as potential delay and 10 were classified as okay using Speed DIAL. When using the Virginia Standards of Learning third-grade total English test scores as the outcome criterion, Speed DIAL's index of sensitivity was 35.9%. Thirty-nine students were classified as failing the Virginia Standards of Learning third-grade total English test. Among those 39 students, 14 were originally classified as potential delay and 25 classified as okay using Speed DIAL.

Index of specificity is the percentage of students who passed the Virginia Standards of Learning third-grade tests and were correctly classified by Speed DIAL as okay four years earlier. When using the Virginia Standards of Learning third-grade math test scores as the outcome criterion, Speed DIAL's index of specificity was 82.8%. One hundred eighty-six students were classified as passing the Virginia Standards of Learning third-grade math test. Among these 186 students, 32 were originally classified as potential delay and 154 were classified as okay using Speed DIAL.

When using the Virginia Standards of Learning third-grade total English test scores as the outcome criterion, Speed DIAL's index of specificity was 84.9%. One hundred sixty-six students were classified as passing the Virginia Standards of Learning third-grade total English test. Among these 166 students, 26 were originally classified as potential delay and 140 were classified as okay using Speed DIAL.

Also of interest are the percentages of correct identifications of potential delay and okay using Speed DIAL. Thirty-eight students were originally classified as potential delay. Six of the 38 students, or 15.8% of those classified as potential delay subsequently were classified as failing the Virginia Standards of Learning third-grade math test. Of the 40 students originally classified by Speed DIAL as potential delay, 14 (35%) subsequently were classified as failing the Virginia Standards of Learning third-grade total English test. Conversely, 164 students were originally classified by Speed DIAL as okay. One hundred fifty-four out of 164 (94%) originally identified as okay using Speed DIAL were subsequently classified as passing the Virginia Standards of Learning third-grade math test. Of the 165 students originally classified as okay using Speed DIAL, 140 (84.8%) subsequently were classified as passing the Virginia Standards of Learning third-grade total English test.

Research Question 3

Is there an association between gender, age, and Speed DIAL scores?

A two-way ANOVA was conducted to evaluate associations between gender, age, and Speed DIAL scores. The gender by age interaction was determined to be not significant. The null hypothesis for gender was rejected. There was an association between gender and Speed DIAL scores, as females scored almost four points higher than did males on Speed DIAL. The null hypothesis for age was retained as there was no association between age and Speed DIAL scores.

Research Question 4

Is there an association between Speed DIAL scores and the intervening variables: retention status, special education status, and number of years of participation in Title I math, Title I reading, and PALS?

All five null hypotheses were rejected, as a negative association was found between each of the five intervening variables and Speed DIAL. The Pearson's correlation analysis showed that lower Speed DIAL scores were associated with a greater likelihood that a student would be retained. Likewise, a student had a greater possibility of being served by special education services if his or her score on Speed DIAL was lower. Finally, lower Speed DIAL scores increased the number of years a student participated in Title I math and reading programs as well as years in the PALS phonological awareness program.

Research Question 5

Is there an association between the Virginia Standards of Learning third-grade math test scores and gender, age, Speed DIAL scores, retention status, special education status, and number of years in the Title I math program?

The null hypotheses for age and retention status were retained. The null hypothesis for gender was rejected with males having higher math scores than females. The null hypothesis for special education status was also rejected with students who had received special education

services having lower math scores than students who did not receive the services. The null hypothesis for Speed DIAL scores was rejected: Students with higher scores on Speed DIAL had higher scores on math. Finally, the null hypothesis for the number of years in Title I math was rejected showing that the longer students had participated in the Title I Math program, the lower the math score.

Research Question 6

Is there an association between the Virginia Standards of Learning third-grade total English test scores and gender, age, Speed DIAL scores, retention status, special education status, number of years in the Title I reading program, and number of years in the Phonological Awareness Literacy Screening (PALS) program?

The null hypotheses for gender, age, Speed DIAL scores, retention status, and years in PALS program were retained. The null hypothesis for special education status was rejected with special education students having lower total English scores than did those students not receiving special education services. Finally, the null hypothesis for the number of years in Title I reading was rejected showing that the longer students had participated in the Title I reading program, the lower the total English score.

Conclusions

Conclusion #1

Speed DIAL does possess some predictive qualities. The analysis using chi-square showed there was a significant association between Speed DIAL classifications of *potential delay* and *okay* and pass-fail classifications from the Virginia Standards of Learning third-grade total English test.

Conclusion #2

Carran and Scott (1992) analyzed eight screening tests that had an average Specificity Index of 91%, a Sensitivity Index of 48% and a positive predictive value of 65%. Gredler (1997) argued that such percentages showed mixed results for the validity of screening instruments. For this study, using the Virginia Standards of Learning third-grade math test as the outcome criterion, Speed DIAL's Sensitivity Index was 37.5%. The Specificity Index was 82.8% and the positive predictive value was 15.8%. These percentages indicate that of all students who were classified as passing by the Virginia Standards of Learning third-grade math test, nearly 83% had been originally identified as okay using the Speed DIAL. However, of all the students who were classified as failing by the Virginia Standards of Learning third-grade math test, only 37.5% had originally been classified as potential delay by using Speed DIAL. The positive predictive value of 15.8% means that only 15.8% of those students originally classified as potential delay using Speed DIAL were later classified as failing the Virginia Standards of Learning third-grade math test.

When using the Virginia Standards of Learning third-grade total English test scores as the outcome criterion, Speed DIAL's Sensitivity Index was 35.9%. The Specificity Index was 84.9%, and the positive predictive value was 35%. These percentages indicate that of all students who were classified as passing by the Virginia Standards of Learning third-grade total English test, nearly 85% had been originally identified as okay using the Speed DIAL. However, of all the students who were classified as failing by the Virginia Standards of Learning third-grade total English test, only 35.9% had originally been classified as potential delay by using Speed DIAL. The positive predictive value of 35% means that only 35% of those students originally classified as potential delay using Speed DIAL were later classified as failing the Virginia Standards of Learning third-grade total English test.

Speed DIAL's Overall Effectiveness Indices were 79% for math scores and 75% when using total English scores. It is also important to look at the implications of the other indices to

determine what they might indicate not only about the worthiness of a screening assessment, but also the individual success of students. While it is true that in this study Speed DIAL possesses poor positive predictive value, the more important issue is one of student outcomes. Consider the Virginia Standards of Learning third-grade math test as an example. First, it should be noted that of the 202 students included in this study who took the assessment, 186 or over 92% of the students were classified as passing. That passing rate is significantly higher than it was eight years ago when the tests were first introduced. Secondly, of the 202 students who took the Virginia Standards of Learning third-grade math test, 38 were originally identified as potential delay using the Speed DIAL. Had Speed DIAL possessed 100% positive predictive value, all 38 of those students would have been classified as failing the Virginia Standards of Learning third-grade math test. Unfortunately for Speed DIAL's positive predictive value, but fortunately for the students themselves, 32 out of the original 38 identified by Speed DIAL as potential delay went on to be classified as passing the Virginia Standards of Learning third-grade math test.

Likewise, 166 out of 205 (81%) students in this study were classified as passing the Virginia Standards of Learning third-grade total English test. Again, much like the math passing rate, the passing rate for total English has increased significantly over the past eight years. Of the 40 students originally identified as potential delay using Speed DIAL, 26 were classified as passing the Standards of Learning third-grade total English test. It could be argued that this was simply a case where the school system inappropriately selected the cut-off criterion resulting in over-referrals or the original designations were correct and the school system's remediation and assistance programs were successful in bridging the gap between performance on the Speed DIAL and Virginia Standards of Learning third-grade tests. A high positive predictive value for the screening instrument would say volumes about what occurred after the screening and before the outcome criteria. While the researcher in this study attempted to document some of these postscreening interventions and accommodations, it was impossible to identify every relevant

variable after the fact. The score from a screening instrument is only relevant if it is subsequently used to make decisions that positively impact future student outcomes.

Conclusion #3

In this study, it was determined through the use of a two-way ANOVA that there was an association between gender and Speed DIAL scores. Females scored higher on Speed DIAL than did males. The mean score for females was 95.3 or almost four points higher than the mean score (91.4) for males. A study by Docherty (1983), mentioned previously in chapter 2, revealed an association between gender and DIAL scores. In that study, *t* tests for gender differences revealed that females scored higher than males did on the DIAL.

Conclusion #4

As noted before, within the study of test predictive validity, there seems to be confusion as to what constitutes a meaningful correlation. A correlation might be noted as high in one study and moderate in another. In one study, a correlation of .35 could be viewed as moderate whereas another study might report that same correlation as low. Even with studies of the same screening instrument, there is a lack of agreement. For example, Ames and Ilg (1964) examined the Gesell School Readiness Test and reported a predictive validity correlation in the moderate range of .74. Other researchers such as Popovics (1982) and Banerji (1992) who also analyzed the Gesell School Readiness Test achieved correlations ranging from .11 to .39.

Pearson's correlation was used to answer Research Question 3 concerning the association between Speed DIAL and the intervening variables. Pearson's correlation can range from -1 to +1. Correlations of zero represent no association and associations are considered stronger as they get closer to -1 or +1. The square of the *r* value is the percentage of variance in the dependent variable that is accounted for by membership in the independent variable groups. According to Cohen's (1988) Standard, each correlation (r) has a corresponding r^2 and the r^2

values correspond to the effect size, with effect sizes of .01 considered small, .06 considered medium, and .14 considered large. As with most research studies, there are many variables that cannot be controlled for, such as home environment. What is considered a small or large correlation is dependent upon the particulars special to each area of investigation. Because of this, any correlation above or below zero can be considered as indicative of an association. In this case, the correlations, while not considered strong, do indicate negative associations between Speed DIAL and the intervention variables.

Conclusion #5

A negative association between Speed DIAL and the intervening variables would be expected. Again, the analysis indicated that lower Speed DIAL scores were associated with greater number of years participating in Title I math services. In fact, 10.9% (r^2 = .109) of the variance in years participating in Title I math was accounted for by Speed DIAL scores. Lower Speed DIAL scores were also associated with greater years of participation in Title I reading and PALS program as well as a greater opportunity of being retained and/or served by special education services. Speed DIAL accounted for 6.3% of the variance in retention status (r^2 = .063), 2.9% of the variance in special education status (r^2 = .029), 7.3% of the variance in years participating in the PALS program (r^2 = .073) and 10.2% of the variance in number of years in Title I reading (r^2 = .102). It is reasonable to expect that students scoring low on the Speed DIAL would therefore be eligible to participate in future interventions. It is also reasonable to expect that the students would spend more time in these interventions than would their higher scoring counterparts.

Conclusion #6

A primary focus of this study was on the association between Speed DIAL and the dependent variables, Virginia Standards of Learning third-grade math and Total English scores.

Using a three-step hierarchical regression model for each dependent variable, I was interested in the association between Speed DIAL and the dependent variable at Step 2 of the regression model and the change, if any, in that association in Step 3 after the intervention variables (special education status, retention status, and other intervention variables) had been included in the model.

In the regression model for math, the *Beta* coefficient for Speed DIAL at Step 2 was .374 (p = .001). However, the coefficient for Speed DIAL at Step 3 was .23 (p = .003). Because the coefficient for Speed DIAL was statistically significant at both Steps 2 and 3, the association between Speed DIAL and math decreased substantially between Steps 2 and 3.

In the regression model for total English, the association between Speed DIAL and total English scores in Step 2 showed the *Beta* coefficient for Speed DIAL was .301 (p = 01). At Step 3, the coefficient for Speed DIAL diminished substantially to .143 and was no longer statistically significant (p =.09).

The interpretation of these findings for Speed DIAL was guided by a model called The Elaboration Paradigm (Babbie, 1989). The Elaboration Paradigm is a logical model used to interpret the findings of what happens to an association between two variables after control variables are introduced into the model. According to Babbie, when an association between two variables diminishes or vanishes after a control variable is introduced, as was the case for Speed DIAL in both regression models, the meaning of this finding is based on the time order of the control variable relative to the independent (Speed DIAL) and dependent variable (either Virginia Standards of Learning third-grade math or total English). When the control variable is an antecedent variable, occurring in time prior to the independent and dependent variable, the original association is spurious. Alternatively, if the control variable is an intervening variable, intervening between the independent and dependent variable, the original association is not spurious. Instead, this is called interpretation whereby the control variables are merely the

mechanisms through which the association between Speed DIAL and the dependent variable occurs (Babbie).

The association between Speed DIAL and the dependent variable in both regression models diminished in Step 3 not because the association was spurious or because Speed DIAL had no effect on the dependent variable. Instead, the coefficient for Speed DIAL in Step 3 diminished substantially for two reasons: (a) Speed DIAL was negatively related to the intervention variables as was found in the analysis of Research Question 4; and (b) the intervention variables intervened between Speed DIAL and the dependent variable. These findings support the conclusion that Speed DIAL has value in predicting the outcome of Virginia Standards of Learning math and Total English scores. The association between Speed DIAL and either Virginia Standards of Learning math or Total English scores is not spurious, but rather supported. The intervention variables were merely the mechanisms through which that association was interpreted.

In summary, the findings in this study support the conclusion that Speed DIAL has some value in predicting the outcome of the Virginia Standards of Learning math and Total English scores.

Recommendations for Further Practice

The following are recommendations for practice:

1. Considering the low positive predictive value that could be interpreted as over-referrals, the school system should consider using the full DIAL-3 and compare the results to those of Speed DIAL.
2. The school system should consider the cut-off level chosen to interpret the results. In this study, a cut-off level of 1 *SD* was chosen by the school officials to classify students as either okay or potential delay based on the students' performance on the screening instrument. Using a cut-off level of 1 *SD* identified over 19.5 % of the

population as potential delay, when ideally 1 *SD* should have identified about 16% of the population based on national standardization of the instrument. One could interpret the low positive predictive values as over-referrals when using math and total English as the outcome criteria. By moving the cut-off from 1 to 1.3 *SD* from the mean, the school system could conceivably lower the number of over-referrals and better serve the remaining children classified as potential delay.

3. As recommended by Dworkin (1989), the school system should view results from the Speed DIAL or any screening instrument with a certain degree of skepticism, and realize that there are many factors and opportunities for error in the screening process. In addition to screening, teachers and administrators should continue to monitor students throughout the school year, which is a continuous process that is more ongoing than simply screening one time.

Recommendations for Further Research

The fact that many schools use screening tests such as the Speed DIAL Developmental Screening Test underscores the need to gather and analyze the available data about the assessments used as predictive indicators. Because screening instruments have far reaching repercussions for the future academic success of a child, it is important that the instruments used, in this case, Speed DIAL possess a high degree of reliability and validity (Gredler, 1997). The following are recommendations for further research:

1. A replication of this study should be conducted in another school system.
2. A replication of this study should be conducted using an outcome criterion other than the Virginia Standards of Learning third-grade tests.
3. A replication of this study is needed using a larger population size and/or analyzing more than two years of data.

4. There should be an implementation of a longitudinal qualitative study evaluating the same group of students from preschool through third grade, whereby the researcher could better document the interventions, accommodations, differentiated instruction, strategies, and remediation provided to students as well as the differences within classrooms and schools.
5. Execution of other studies is needed using other screening instruments, such as the DIAL-3 and their predictive validity based on the Virginia Standards of learning third-grade tests.
6. Implementation of a longitudinal study is needed evaluating the same students who participated in this study as they progress and use later performance in the upper grades as the outcome criteria from which to measure the predictive validity of Speed DIAL.

REFERENCES

Action Team on School Readiness. (1992). *Every child ready for school*. Washington, DC: National Governor's Association.

Agostin, T. M., & Bain, S. K. (1997). Predicting early school success with developmental and social skills screeners. *Psychology in the Schools*, *34,* 219-228.

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

Ausubel, D. P., Sullivan, E. V., & Ives, S. W. (1980). *Theory and problems of child development* (3rd ed.). New York: Columbia University Press.

Babbie, E. R. (1989). *The practice of social research* (5th ed.). Belmont, CA: Wadsworth.

Banerji, M. (1992). An integrated study of the predictive properties of the Gesell School Readiness Screening Test. *Journal of Psychoeducational Assessment*, *10,* 240-256.

Book, R. M. (1980). Identification of educationally at-risk children during the kindergarten year: A four-year follow-up study of group test performance. *Psychology in the Schools*, *17,* 153-158.

Bracken, B. A. (1987). Limitations of preschool instruments and standards for minimal levels of technical adequacy. *Journal of Psychoeducational Assessment*, *4,* 373-326.

Bracken, B. A. (1992). *The psychoeducational assessment of preschool children*. Needham Heights, MA: Allyn & Bacon.

Bredekamp, S., & Rosegrant, T. (1992). *Reaching potentials: Appropriate curriculum and assessment for young children* (Vol. 1). Washington, DC: National Association of the Education of Young Children.

Bredekamp, S., & Shepard, L. (1989). How best to protect children from inappropriate school expectations, practices, and policies. *Young Children*, *44,* 14-24.

Carran, D. T., & Scott, K. G. (1992). Risk assessment in preschool children: Research implications for the early detection of educational handicaps. *Topics in Early Childhood Special Education, 12,* 196-211.

Chew, A. L., & Lang, W. S. (1990). Predicting academic achievement in kindergarten and first grade drop prekindergarten scores on the Lollipop test and DIAL. *Educational and Psychological Measurement*, *50,* 431-437.

Clark, A. M. (1982). Developmental discontinuities: An approach to assessing their nature. In L.A. Bond & J. M. Joffe (Eds.), *Facilitating infant and early childhood development* (pp. 58-77). Hanover, NH: University Press of New England.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum.

Costenbader, V., Rohrer, A. M., & Difonzo, N. (2000). Kindergarten screening: A survey on current practice. *Psychology in the Schools*, *37,* 323-332.

Czudnowski, C., & Goldenberg, D. (1998). *Examiner's manual: DIAL-3, developmental indicators for the assessment of learning*. Circle Pines, MN: American Guidance Service.

DeCos, P. L. (1997). *Readiness for kindergarten: What does it mean?* Sacramento, CA: California Research Bureau, California State Library.

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

Docherty, E. M. J. (1983). The DIAL: Preschool screening for learning problems. *The Journal of Special Education*, *17,* 195-202.

Drieling, L., & Copeland, E. (1988). *A model for the evaluation of kindergarten screening tests*. Paper presented at the Annual Convention of the American Psychological Association, Atlanta.

Durkin, D. (1987). Testing in the kindergarten. *The Reading Teacher*, *40,* 766-770.

Dworkin, P. H. (1989). Developmental screening: Expecting the impossible? *Pediatrics, 83,* 619-622.

Dziuban, C., & Mealor, D. (1982) Validity of a primary screening device as a predictor of subsequent academic achievement. *Perceptual and Motor Skills*, *54,* 1053-1054.

Eaves, L. C., Kendall, D. C., & Crichton, J. U. (1974). The early identification of learning disabilities: A follow-up study. *Journal of Learning Disabilities*, *7,* 632-638.

Frankenburg, W. K. (1985). The concept of screening revisited. In W. F. Frankenburg, R. W. Ende, & J. W. Sullivan (Eds.), *Early identification of children at-risk* (pp. 3-17). New York: Plenum Press.

Funk, S. G., Sturner, R. A., & Green, J. A. (1986). Preschool prediction of early school performance: Relationship of McCarthy scales of children's abilities prior to school entry to achievement in kindergarten, first, and second grades. *Journal of School Psychology*, *24,* 181-194.

Gagne, R. M. (1985). *The conditions of learning* (4th ed.). New York: Holt, Rinehart, & Winston.

Gall, M. D., Borg, W. R., & Gall, J. P. (1996). *Educational research: An introduction*. New York: Longman.

Gallerani, D., O'Regan, M., & Reinherz, H. (1982). Prekindergarten screening: How well does it predict readiness for first grade? *Psychology in the Schools*, *19,* 175-182.

Gesell, A. (1940). *The first five years of life*. New York: Harper & Brothers.

Goodenough, F. (1949). *Mental testing.* New York: Rinehart.

Gredler, G. R. (1976). Issues in school readiness. In G. Bea., K. Minke, & A. Thomas (Eds.), *Children's needs II: Developmental problems and alternatives* (pp. 489-499). Bethesda, MD: National Association of School Psychologists.

Gredler, G. R. (1992). *School readiness: Assessment and educational issues*. Brandon, VT: Clinical Psychology Publishing.

Gredler, G. R. (1997). Issues in early childhood screening and assessment. *Psychology in the Schools*, *34,* 99-106.

Heward, W. E. (1996). *Exceptional children: An introduction to special education*. Englewood Cliffs, NJ: Prentice-Hall.

Hinkle, D. E., Wiersma,W., & Jurs, S. G. (1998). *Applied statistics for the behavioral sciences* (4th ed.). Boston: Morghton Mifflin.

Hobbs, N. (1975). *The futures of children*. San Francisco: Josey-Bass.

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

Hunt, J. M. (1969). The impact and limitations of the giant of developmental psychology. In D. Elkind & J. Flavell (Eds.), *Studies in cognitive development: Essays in honor of Jean Piaget* (42-43). New York: Oxford University Press.

Jackson, G. B. (1975). The research evidence on the effects of grade retention. *Review of Educational Research*, 45, 613-635.

Kaplan, C. (1996). Predictive validity of the WPPSI-R: A four year follow-up study. *Psychology in the Schools*, *33,* 211-220.

Klein, A. E. (1977). The validity of the screening test of academic readiness in predicting achievement in the first and second grades. *Educational and Psychological Measurement*, *37,* 493-499.

Kurdek, L. A., & Sinclair, R. J. (2001). Predicting reading and mathematic achievement in fourth-grade children from kindergarten readiness scores. *Journal of Educational Psychology*, *93,* 451-455.

Lazar, I., & Darlington, R. (1982). Lasting effects of early education: A report from the consortium for longitudinal studies. *Monographs of the Society for Research in Child Learning Disabilities, 195,* 47.

Leach, D. J. (1980). Considerations in screening for school learning difficulties: An overview of practice. *Educational Studies*, *6,* 181-197.

Lichtenstein, R., & Ireton, H. (1984). *Preschool screening, identifying young children with developmental and educational problems*. United Kingdom: Grune & Stratton.

Mallory, B. L., & Kerns, G. M. (1988). Consequences of categorical labeling of preschool children. *Topics in Early Childhood Education*, *8,* 39-50.

Mantzicopoulos, P. Y. (2000). Can the brigance k & 1 screen detect cognitive/academic giftedness when used with preschoolers from economically disadvantaged backgrounds? *Roeper Review*, *22,* 185-191.

Meisels, S. J. (1976). *Developmental screening: A guide*. Washington, DC: National Association for the Young Child.

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

Meisels, S. J. (1987a). High stakes testing in kindergarten. *Educational Leadership*, *46,* 16-22.

Meisels, S. J. (1987b). Uses and abused of developmental screening and school readiness testing. *Young Children*, *42,* 68-73.

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

Messick, S. (1980). Test validity and the ethics of assessment. *American Psychologist*, *35,* 1012-1027.

Neill, D., & Medina, N. (1989). Standardized testing: Harmful to educational health. *Phi Delta Kappan*, *70,* 688-697.

Obrzut, J. E., Bolocofsky, D. N., Heath, C. P., & Jones, M. J. (1981). An investigation of the DIAL as a pre-kindergarten screening instrument. *Education and Psychological Measurement*, *41,* 1231-1241.

Paget, K., & Bracken, B. (1983). *The psychoeducational assessment of preschool children.* New York: Grune & Stratton.

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

Phonological Awareness Literacy Screening Technical Reference. (2004). Retrieved March 13, 2005, from http://www.pals.virginia.edu/

Piaget, J. (1970). *Genetic epistemology*. (E. Duckworth, Trans.). New York: Columbia University Press.

Popovics, A. J. (1982). Total and component actuarial scores of the Gesell copy form test as predictors of achievement, intelligence, and creativity measures. *Journal of Psychology*, *110,* 293-295.

Potton, A. (1983). *Screening*. London: MacMillan Education.

Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom*. New York: Holt, Rinehart & Winston.

Salvia, J., & Ysseldyke, J. (1998). *Assessment* (7th ed.). Boston: Houghton Mifflin.

Schmidt, S., & Perino, J. (1985). Kindergarten screening results as predictors of academic achievement, potential, and placement in second grade. *Psychology in the Schools*, *22,* 146-151.

Shepard, L.A. (1997). Children not ready to learn? The invalidity of school readiness testing. *Psychology in the Schools*, *34,* 85-97.

Shepard, L.A., & Smith, M. L. (1986). Synthesis of research on school readiness and kindergarten retention. *Educational Leadership*, *44,* 78-86.

Simmer, M. L., & Barnes, M. J. (1991). Relationships between first-grade marks and the high school dropout problem. *Journal of School Psychology*, *29,* 331-335.

Southworth, L. E., Burr, R. L., & Cox, A. E. (1980). *Screening and evaluating the young child*. Springfield, IL: Charles C. Thomas.

Steinbauer, E., & Heller, M. S. (1978). The Boehm test of basic concepts as a predictor of academic achievement in grade 2. *Psychology in the Schools*, *15,* 357-360.

Stone, B. J., & Gridley, B. E. (1991). Test bias of a kindergarten screening battery: Predicting achievement for white and native American elementary students. *School Psychology Review*, *20,* 132-139.

Stott, L. H., & Ball, R. S. (1965). Infant and preschool mental tests: Review and evaluation. *Monographs of the Society for Research in Child Development*, *30,* 3.

Suen, H. K., Czudnowski, G. M., & Goldenberg, D. S. (1989). Classification reliability of the DIAL-R preschool screening test. *Educational and Psychological Measurement, 49,* 673-680.

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

United States Department of Education. (2003) *Condition of education 2003*. Washington, DC: Author.

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

Virginia Department of Education. (1999). *Standards of learning assessments (SOL) tests validity and reliability information: Spring 1998 administration report*. Richmond, VA: Author.

Virginia Department of Education. (2000). *Virginia standards of learning assessments: Technical report.* Richmond, VA: Author.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Wenner, G. (1995). Kindergarten screens as tools for the early identification of children at risk for remediation or grade retention. *Psychology in the Schools*, *32,* 249-54.

APPENDIX

Letter to Superintendent of Schools

10824 #### ####### Highway
#####, ###### #####
February 4, 2005

##.####### ######
Superintendent, #### ###### Schools
P.O. Box ####
####, ###### #####

Dear ##. ######:

I am interested in the predictive validity of the pre-kindergarten screening instrument and would like to investigate the correlation between student scores on the Speed DIAL and Scores from SOL tests. I will be using scores from the years 1999 through 2005. I will record the age, gender of the student, participation in interventions, and consider these factors in the data analysis. Finding an association between these scores could aid school personnel in making better decisions about screening instruments, student placements, and instructional choices.

No identifying data will be recorded in the process of gathering student scores, and #### ##### Schools will not be referenced in this study. ##. #### ###### and ##. ##### ##### have graciously agreed to serve as liaisons in the data collection process to assure the confidentiality of students and the school system. I will be conducting this study under the supervision of East Tennessee State University.

I believe that the results of this study will enable us to better serve the students of #### #####. I appreciate your consideration of this matter.

Sincerely,

Roger Walk

Roger Walk has my permission to conduct the above mentioned study within the #### ##### School Division

_____
Superintendent, #### ###### Schools

VITA

ROGER ANTHONY WALK

Personal Data:	Date of Birth: April 1, 1971 Place of Birth: Norton, Virginia Marital Status: Married
Education:	University of Virginia's College at Wise Bachelor of Arts 1994
	East Tennessee State University, Johnson City, Tennessee; Educational Leadership and Policy Analysis, M.Ed.; 1999
	East Tennessee State University, Johnson City, Tennessee; Educational Leadership and Policy Analysis, Ed.D.; 2005
Professional Experience:	Title I Reading/Math Teacher, Coeburn Primary School; Wise County Schools; 1996-1997
	Kindergarten Teacher, Appalachia Elementary School; Wise County Schools; 1997-1998
	First Grade Teacher, Appalachia Elementary School; Wise County Schools; 1998-1999
	Second Grade Teacher, Appalachia Elementary School; Wise County Schools; 1999-2000
	Third Grade Teacher, Appalachia Elementary School; Wise County Schools; 2000- 2001

Professional
Experience:

First Grade Teacher, Appalachia Elementary School;
Wise County Schools;
2001- 2002

Third Grade Teacher, Appalachia Elementary School;
Wise County Schools;
2002–2003

Director, 21st Community Learning Center;
J.W. Adams Combined School;
Wise County Schools;
2003-2005

Principal, Mountain View Elementary School;
Johnson City Schools;
2005-present