A comparison of the subtests of the Gesell School Readiness Screening Test as predictors of reading achievement

1984

Dolores W. Sapp
University of Central Florida

Find similar works at: https://stars.library.ucf.edu/rtd

University of Central Florida Libraries http://library.ucf.edu

Part of the Educational Administration and Supervision Commons

STARS Citation


This Doctoral Dissertation (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of STARS. For more information, please contact lee.dotson@ucf.edu.
A COMPARISON OF THE SUBTESTS
OF THE GESELL SCHOOL READINESS SCREENING TEST
AS PREDICTORS OF READING ACHIEVEMENT

by

DOLORES W. SAPP

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Education from the Department of Administration Supervision in the College of Education University of Central Florida Orlando, Florida

December 1984

Major Professor: Dr. Patricia Manning
I should like to acknowledge the particular assistance given to me by the members of my committee at the University of Central Florida; Dr. Patricia C. Manning, Chairman, Dr. Robert R. Lange, Dr. Richard D. Tucker, and Dr. Elaine B. Cox. This undertaking would not have been completed with such care to detail and accuracy without their assistance. Special appreciation is due the School Board of Volusia County and the many school people who gave their assistance to this study. I must especially thank my husband, Jerry, for his patience during this entire project.
# TABLE OF CONTENTS

LIST OF TABLES ........................................ iii

CHAPTER 1 THE PROBLEM .................................. 1
  Statement of the Problem ............................... 4
  Study Questions ....................................... 4
  Definition of Terms ................................... 4
  Limitations of the Study ............................. 5
  Assumptions .......................................... 6

CHAPTER 2 REVIEW OF THE LITERATURE ................. 7
  Readiness for School ................................ 8
  Predictors of School Success ......................... 16
  Success Following an Additional Year of Kindergarten 25
  Gesell Subtests as Predictors ....................... 28

CHAPTER 3 METHODS AND PROCEDURES ................. 30
  Population and Sample Selection .................... 30
  Instruments .......................................... 32
  Validity and Reliability ............................ 34
  Data Collection ..................................... 36
  Research Design and Treatment ...................... 37

CHAPTER 4 ANALYSIS OF DATA ......................... 39
  Correlation .......................................... 39
  Multiple Regression ................................ 45
  Descriptive Review .................................. 56

CHAPTER 5 SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS ............. 69
  Summary and Conclusions ............................ 70
  Recommendations ..................................... 79

APPENDIX ............................................... 81

REFERENCES AND NOTES ................................ 114
LIST OF TABLES

1. Intercorrelations Between Independent Variables . 40
2. Correlations Between Predictor and Dependent Variables ............... 43
3. Summary of Stepwise Regression for Word Attack: Gesell Subtests Entered First ........... 48
4. Summary of Stepwise Regression for Vocabulary: Gesell Subtests Entered First ........... 51
5. Summary of Stepwise Regression for Reading Comprehension: Gesell Subtests Entered First . 53
6. Summary of Stepwise Regression for Language: Gesell Subtests Entered First ........... 55
7. Comparison of Achievement Test Section for Each Developmental Age of the Gesell Cubes Test .. 58
8. Comparison of Achievement Test Section for Each Developmental Age of the Gesell Name Test ... 59
9. Comparison of Achievement Test Section for Each Developmental Age of the Gesell Copy Forms Test 60
10. Comparison of Achievement Test Section for Each Developmental Age of the Gesell Numbers Test . 61
11. Comparison of Achievement Test Section for Each Developmental Age of the Gesell Incomplete Man Test ............... 62
12. Comparison of Achievement Test Section for Each Developmental Age of the Gesell Interview Test . 63
13. Comparison of Achievement Test Section for Each Developmental Age of the Gesell Animals Test . 64
14. Comparison of Achievement Test Section for Each Developmental Age of the Gesell Interests Test . 65
A COMPARISON OF THE SUBTESTS
OF THE GESELL SCHOOL READINESS SCREENING TEST
AS PREDICTORS OF READING ACHIEVEMENT

by

DOLORES W. SAPP

A dissertation submitted in partial fulfillment of the requirements for the degree of
DOCTOR OF EDUCATION in the Department of ADMINISTRATION AND SUPERVISION at
the University of Central Florida
Orlando, Florida

December 1984

Major Professor: Dr. Patricia C. Manning

ABSTRACT

The major purpose of this study was to determine which subtests of the Gesell School Readiness Screening Test are the best predictors of reading achievement. The procedure used was a stepwise multiple regression to determine which subtests were the best predictors of achievement for each reading section of the Comprehensive Test of Basic Skills.

Seventy-two students from four schools were evaluated for this study. Three testing instruments were administered
to these students. They were the Gesell test, the Otis-Lennon test, and the reading subtests of the Comprehensive Test of Basic Skills.

The Gesell test consisted of eight subtests: Cubes, Name, Copy Forms, Numbers, Incomplete Man, Interview, Animals, and Interests. The best single predictor subtest was Interview.

Four separate regression analyses were computed, one for each of the subtests of the reading section of the Comprehensive Test of Basic Skills. In each case, the Gesell subtest Interview entered the stepwise regression at step one. After the information for Interview was entered into the regression equation, the remaining seven subtest scores added very little to the prediction.

In summary, although there was some predictive ability in the Gesell test, the level of predictability was not strong. Post factum interpretation of the low predictability was presented as were recommendations for Volusia County in the use of the Gesell test for placement of students in the kindergarten-first grade transitional class.
CHAPTER 1
THE PROBLEM

The Gesell School Readiness Test is a readiness test designed for administration individually to children from the ages of four to eight. The test is divided into eight subtests, each consisting of a series of tasks. The subtests are: Cubes, Name, Copy Forms, Writing Numbers, Incomplete Man, Initial Interview, Animals, and Interests.

For over a decade many educators in Volusia County have been concerned about children who were entering the school academic arena of first grade arbitrarily because they would be six years old on or before January first of the following year. Florida students began school at a comparatively young age. The Florida Educational Accountability Act of 1976, which included the Primary Education Program, was passed and included a provision for the rollback of the chronological age of students entering school. However, almost 20 percent of the students continued to have difficulty succeeding in school. During the 1982-83 school year, primary resource teachers in all Volusia County elementary schools were trained and certified as Gesell School Readiness Screening Test administrators. The process was an effort to help
educators become cognizant of the developmental differences among children. The teachers were trained to administer the subtests that pertained to children under seven years of age.

During that school year (1982-83), over 60 percent of the kindergarten students in Volusia County were tested with the Gesell test to determine their developmental age. A decision on the number of students to be tested at each school center was made at the school site. Tests that determined developmental ages were used to assist in placement of these students for the 1983-84 school year in kindergarten, transitional kindergarten-first grade, or first grade.

The Gesell School Readiness Screening Test takes from 30 to 45 minutes to administer; an equal amount of time is required to score or interpret the behaviors recorded during the test. If some subtests of the test are found to be high predictors of reading achievement, part of the test could be eliminated. The time saved could be used to test more students and achieve a more uniform decisionmaking policy.

For considerably more than 50 years, the Gesell Institute has been primarily concerned with outlining and describing the patterned and predictable stages in the development of early human behavior. Much of the work has
focused on developmental readiness and relating readiness to school success. For example, Ilg and Ames (1964) reported that as many as 50 percent of the children who enter the first grade may not be ready for early schooling. There are vast numbers of students, including highly intelligent children, who have been overplaced in school. Even though a child's IQ is high, he still may not have the developmental maturity required for school. The Gesell developmental position is that training and preschool experiences do not increase the rate of mental, psychomotor, or affective growth.

Although current findings pinpoint developmental age as the best predictor of school success, chronological age continues to be the single prerequisite for school entrance. The importance of having children ready for beginning a given grade should not be underestimated. The Gesell position is that children should start school and be subsequently promoted on the basis of their developmental age rather than their chronological age. Gesell (1977) suggested that perhaps 50 percent of school problems could be prevented or remedied by correct placement.

Heimgartner (1970) found that the Gesell School Readiness Screening Test was a good predictor of reading success (determined by student's reading level).
Heimgartner's study, completed in the late 1960s, consisted of 104 students with a mean IQ of 114 attending the University of Northern Colorado Laboratory School and placed in grade levels kindergarten to third grade. He determined that the two subtests, Copy Forms and Incomplete Man, gave the best prediction of developmental age and reading success.

Statement of the Problem

The purpose of this study was to determine which subtests of the Gesell School Readiness Screening Test best predicted reading achievement.

Study Questions

1. To what degree do the subtests of the Gesell School Readiness Screening Test at kindergarten level predict Word Attack, Vocabulary, Reading Comprehension, and Language scores on a standardized first grade achievement test (Comprehensive Test of Basic Skills)?

2. Which combination of subtests predicts reading achievement as well as or nearly as well as the full set of tests?

Definition of Terms

For the purpose of this study the definitions of key terms are as follows:

Developmental Age. The age score a child received for each
subtest of the Gesell School Readiness Screening Test.

Intelligence Quotient. An IQ score obtained from the Otis-Lennon School Abilities Test.

Gesell School Readiness Screening Test. A readiness test composed of eight subtests, each consisting of a series of tasks.

Otis-Lennon School Abilities Test. A group test designed to provide a measure of general intellectual ability.

Primary Education Program. A comprehensive, prescriptive program of primary education in kindergarten through third grade.

Primary Resource Teacher. A resource teacher assigned to each school center who assists with the implementation of the kindergarten through third grade program.

Reading Achievement. The reading behavior measured by the Comprehensive Test of Basic Skills reported as a scale score.

Limitations of the Study

The study was limited in the following ways:

1. The population of this study was limited to students enrolled in four public schools in Volusia County.

2. One school did not test all of their kindergarten students.

3. Children moving in and out of the school district may have caused the loss of some subjects.
4. One school dropped out of the study in March 1984.

**Assumptions**

1. All of the primary resource teachers involved in the study used appropriate test administration procedures.

2. The student attrition rate did not severely affect the results of the study.

3. The subjects used in this study were representative of the students of Volusia County.

4. The subjects used in this study were representative of the students used in the standardization studies completed by the Comprehensive Test of Basic Skills and the Otis-Lennon School Abilities Test.
CHAPTER 2
REVIEW OF THE LITERATURE

The traditional role and concept of the American child has altered dramatically. This phenomenon can be attributed to the surmounting pressures of an often transient populace which includes working mothers, day-care centers, and fast-food restaurants. Based on his review of the literature, Elkind (1981) stated that economic pressures on the family unit and disintegration of that family unit are forcing the young child to go from a secure, loving, unstructured atmosphere, embodied in the home, to the more regimented environment found in preschools and day-care centers. Thus boys and girls of the 80s are expected and often pressured to achieve academic goals at an earlier chronological age, which often runs counter to their maturational needs. These children are being deprived of the gift of time. Parents, who often attempt to realize their own expectations via their offspring, add to the pressures already placed upon the child by "academia." The stress resulting from this volatile combination of earlier-demanded achievement and a fluctuating, unstructured, rapid-paced society has wrought havoc upon many of our young.
Pressure to improve both the quality and quantity of academic output descends from federal policy to state, county, and local school administrators, and finally to the classroom teacher. Unfortunately the accountability demanded by this infrastructure ignores the emotional, physical, and maturational well-being of the individual child. An unflattering analogy can be made between a factory that mass produces a product with schools that mass produce students "filled to the top" with knowledge deemed academically necessary. Elkind concluded, based on findings as cited by Gesell (1977) and Piaget (1964), that the important concepts of readiness, learning rates, ways of learning, mental abilities, and emotional deficiencies are deemphasized by the overwhelming need to excel.

Readiness for School

Research compiled by a number of educators suggests many stressful experiences that face children upon entry into school would and could be avoided. Brenner (1957), while emphasizing the importance of the interaction of all readiness factors upon successful school achievement, found that age also played a significant role in determining readiness. He stated that parents should hold back younger children. Practical experience in school with children admitted earlier would prove the advisability of delaying entry in school. He referred to the acceptance of public
opinion as a vital force for implementing the readiness concept for school entrance.

In a more recent work, Brenner and Scott (1971) cited the factors necessary in the consideration of readiness for school (physical, mental, emotional, social). Defining readiness for school as a functional interrelationship between an individual child and the demands of the school, Brenner and Scott viewed readiness as a matter of total development, the result of 69 interacting factors from all areas of the personality. They indicated the necessity that the child be able to spend the day without the constant help and comfort of the mother (fostering independence). Thus, Brenner and Scott pointed to the necessity for the "readiness" of the parent to "let go."

In an in-depth study of kindergartners in Central Bucks County, Pennsylvania, Andrews (1971) strongly supported the Gesell School Readiness Screening Test as an indicator of school readiness, with developmental age a consistent measure of predicting success in kindergarten. She found that chronological age and/or IQ alone were not good criteria for admission to kindergarten. Gesell testing embraced all the individual, physical, mental, emotional, and social factors. She further concluded that a child not found to be developmentally ready would likely have a difficult time in the beginning school experience.
Andrews (1971) confirmed that the score on the Gesell test was significantly related to the promotion-retention rate in her study. When correlated by year, developmental age and promotion-retention had a consistently high and significant relationship. When recommendation for delayed entrance was the outcome of Gesell testing, Andrews ascertained that the child was prevented from having a failed experience and the need for remedial services in later grades was often reduced. The benefit derived from determining each child's developmental age was in providing an appropriate learning situation for them. Thus, she concluded that entrance to kindergarten should be flexible and realistic with entrance based upon a minimum developmental age rather than a set chronological age.

Ogletree (1972) and Piaget and Inhelder (1964) concurred with Brenner and Scott (1971) and Andrews (1971) in a holistic view of child development as it pertains to readiness for school. They espoused the theory that a child gradually gains control of the locomotive, speech, and cognitive processes (thinking being a subtle form of movement). They maintained that the orderly sequence of neuromatomatic growth is such that the capabilities for certain learning performances fall along an age-readiness cycle. Any attempt to speed up cognitive development interferes with physical development, possibly causing a reduction in
full development. Ogletree felt, too, that environmental deprivation retards normal maturation.

Carll and Richard (1977) made the point that school readiness is the ability to learn and cope with the school environment without undue stress and to sustain this level of learning and coping. Readiness to read is not school readiness, although the two are often considered to be synonymous. Many children who are ready to read are not ready for the school situation. Carll and Richard defined overplacement as a child being in over his head. The grade level imposes pressures which inhibit growth in some areas. The overplaced child can be identified as manifesting one or some of the following behaviors: avoidance, conformity, passive resistance, overdrive. They also defined a readiness program as being for those who have faith in the orderliness of nature and who trust a child to know his own needs.

Satz and Friel (1974) found that longitudinal studies by Bloom (1964) and themselves on the growth of primary mental abilities have already shown a sequential and differential rate of development with age. Perceptual and spatial abilities revealed an earlier ontogenetic development, whereas verbal cognitive abilities revealed a later and slower development with age. This theory was compatible with those of Piaget (1964) and Bruner (1968),
who postulated that the child goes through consecutive stages of thought during development. Each stage incorporated the processes of the preceding stage into a more complex and hierarchically integrated form of adaptation. Satz and Friel stated that their results strongly suggested that the reading achievement levels of children at the end of first grade could be validly predicted from an assessment of their developmental and neuropsychological performance at the beginning of kindergarten. The results provided an early warning system which detected problems before the child began formal reading at a time when his central nervous system may become more plastic and subject to the shattering effects of repeated academic failure.

Shapiro (1972) saw the school as responsible for fostering the child's psychological development in a broad sense, encompassing affective and social as well as cognitive development. This developmental-interaction approach includes acquiring and ordering information, judging, reasoning, problem-solving, and using systems of symbols, as part of the cognitive functions and the development of self-esteem, a sense of identity, the internalization of impulse control, the capacity for autonomous response, and relatedness to other people as part of the personal and interpersonal process. The basic
tent of the developmental-interaction approach, then, is that growth of cognitive functions cannot be separated from the growth of personal and interpersonal processes. The interdependence of these developmental processes was the *sine qua non* of the developmental-interaction approach. Shapiro stressed the integration rather than the compartmentalization of functions. He further stated that it was a goal of school to minimize the gap between capacity and performance by providing an environment that allowed and encouraged children to do what they are capable of accomplishing.

Citing Biber and Franklin (1967), Shapiro and Biber (1972) averred that progress from earlier to later levels of functioning in any domain (emotional, intellectual, social) is characterized by moments of equilibrium in which the individual's schemata are adequate for the task at hand and by moments of instability in which currently operative structures are breaking down, but new ones are not sufficiently developed to take over completely.

According to Shapiro and Biber, a major part of education was to provide a balance between experiences that help consolidate a child's understanding and those that provide desirable growth-inducing challenges. A child developed from active, physical, and body-centered modes to primarily perceptual and conceptual ones. The crucial
task of the young child was the construction of a sense of himself. The ingredients of self-concept were seen as determining the quality of his encounter with other people, objects, and life in general.

Heffernan (1962), in another study, said that readiness for any learning meant that the child was able to bring to the task abilities equal to the specific situation. The child had to have all these required abilities, not some of them. The lack of any one of them resulted in the child's failure to read regardless of the excellence of the teacher's instructional methods. Heffernan claimed that the overwhelming amount of research evidence of the 20th century vigorously opposed forcing formal instruction upon the child at an early age.

Among those opposed to early academics were DeHirsh, Jansky, and Langford (1966) who stressed the importance of respecting a child's developmental rhythm. Their research concluded that there was a close link between maturational status in kindergarten and later reading and spelling achievement. Furthermore, non-ready children, rather than catching up, often fell further behind.

Weiss (1962) stated that children learn with more zest and in less time if they are more mature. Her findings also indicated that if a child is exposed to a situation before he is ready or capable of making sufficient progress
for self-satisfaction, he developed unfavorable attitudes towards this situation. Early entrants were at a disadvantage, socially and in leadership opportunities, even with a higher intelligence quotient.

According to Jensen (1969), many school problems could be circumvented if more attention were paid to readiness in the primary grades when children's learning is more easily turned off through extinction due to inadequate readiness. The risks of delaying instruction too long seem much less than the possible disadvantages of forcing instruction on a child who is still far from his optimal reading level. Elkind (1970) concluded that it has become apparent that the results of formal instructional programs at both the preschool and the primary level earlier than the ages of six or seven are only temporary and possibly confounding.

Studies by Moore (1979) and Crain (1974) concurred with Elkind and Jensen. Moore maintained that early schooling and taking the young child out of a familial environment are interfering with and endangering the crucial facets of growth in children. Crain held that there is much to be gained from the developmental age position advocated by Gesell. He observed that parents are in a hurry to control, direct, and instruct children rather than watch, enjoy, and appreciate them while giving them a chance to do their own growing.
Kaufman and Kaufman (1972) compared the developmental tasks and criteria of Piaget and Gesell. Their empirical assessment demonstrated a close agreement between the two, supporting the Gesell School Readiness Screening Test as an appropriate measure of behavioral maturity. Additional studies by Carr (1974) concluded that the Gesell test, administered to kindergarten students, could validly predict a proneness to visual-motor problems.

Predictors of School Success

The necessity for a reliable instrument as well as a suitable program of implementation to predict success in school is an issue that is at the forefront of educational consideration. Wood (1981), in his study, "Predicting School Readiness: The Validity of Developmental Age," concluded that cutoff dates for kindergarten and first grade entrance follow no meaningful pattern. He commended the Gesell Developmental Screening Test as an effective instrument for predicting success or failure in kindergarten. Using the Gesell test, Wood demonstrated that the chronological age of children entering kindergarten within a range of four to six years of age was unrelated to eventual success or failure in school. Correct developmental placement would result in a third to a half of all chronologically eligible kindergarten students being recommended as developmentally unready for
kindergarten. He also recommended, as does the Gesell Institute, that the exact critical age for placement in kindergarten should be calculated locally because average developmental ages of children, as well as the developmental level of kindergarten curricula, vary across school districts.

He further stated that the clear evidence of the validity of the Gesell Test contributed important documentation in the readiness field. Children who are mislabeled as having "special needs" simply because of developmental youngness risk a school life of unnecessary failure. By demonstrating the validity of the instrument, Wood substantiated the values of the Gesell measures in developmental assessment and dispelled some of the criticism surrounding them. According to Wood, careful attention to and understanding of developmental differences in children of the same chronological age was one of the best ways to guarantee a successful school career for each child.

From a survey given to fourth, fifth, and sixth graders, Grant (1981) sought to determine a relationship between the developmental level at school entrance and subsequent school success. All students had been given the Gesell Developmental Examination prior to entering first grade. Placement recommendations from the Gesell test were
compared to the need for retentions, special services, and school attitudes as each child progressed. The findings indicated a need to establish readiness programs for six-year-olds whose needs were not being met within a regular first grade program. As a result, Grant recommended that every effort should be made to place students according to developmental maturity rather than automatic placement by a child's chronological age.

Rohwer (1971) advanced the theory that attempts to increase the autonomous cognitive competence of students are more likely to succeed if they are delayed. The approach agreed, in many respects, with the maturational positions inherent in the ideas of Piaget. Citing Elkind, Rohwer postulated that not only is there no clear-cut longitudinal data to support the claims of the lasting effects of preschool instruction, but that there is evidence in the opposite direction. He indicated a negative correlation between early physical maturation and later intellectual attainments. Thus, the longer formal instruction was delayed, up to certain limits, the greater the period of plasticity and the higher the ultimate level of achievement.

Rohwer also stated that the character of learning and thinking processes in adulthood is discontinuous, in a practical sense, with the character of such processes in
early childhood. He indicated that the character of the demands made by intellectual tasks in adulthood is discontinuous with the character of task demands in early childhood. Furthermore, the effects of early school entrance on attitudes toward school were deleterious. The guiding principle of early education should be to provide the child with repeated experiences of gratification resulting from intellectual activity. The evaluation of student progress and of the efficacy of instruction should be in terms of the performance of skills rather than in terms of retention of skills.

Strom (1965) wrote that the public seems to be unaware of the fact that for pupils who are less mature, formal learning at too early an age may be frustrating by presenting challenges that are too great. Early formal training does not raise the child's learning potential and may even drain the child physically, especially if readiness is taught too early. The question, according to Strom, is: What type of education is most desirable? Efforts in formal education tend to be so structured that they eliminate certain kinds of potential while ostensibly developing readiness in others. Referring to Torrance (1965), Strom continued by indicating one of the major tasks in educating the young is to keep fantasy alive until the child's intellectual development is such that he can
engage in a sound type of creative thinking. Strom states that the trouble is not children eager to give up their fantasy, but adults who insist that children be realistic, stop imagining, and grow up. Besides the fact that pressures for early learning may be injurious to the growth patterns of some children, the recent shift in emphasis from the child's development to his achievement has been harmful. To attach excessive worth to academic achievement could be damaging to personal development. For the child, it could lead to fundamental values that depend not on his character, his personality, or his relations with others, but solely on his academic prowess.

Yonemura (1971) claimed that early emphasis on reading may even violate a child's rights. She suggested a new level of educational interpretation to show children are more likely to do better in school and in life if their minds are allowed to develop in early years.

In a review of research through 1976, Hedges (1976) restated the premise that the child will have more difficulty if the entry age is lowered than he would have at a somewhat older age. The factor of IQ, by itself, is no guarantee that the child is ready socially or emotionally. He concluded that early intensive stimulation is essentially fruitless as it disregards or is not in tandem with the developmental stages of the individual
A state of equilibrium should exist in which social, emotional, and chronological maturities should not be widely different. No child should be placed at a difficulty level in the school program in terms of any one of his developmental capacities. To overlook any one of these factors in the delineation of learning opportunities could result in loss of interest, loss of self-respect, fears, and even the loss of respect from classmates. Hedges maintained that all children, if given time and opportunity, will learn and grow in relation to their "total readiness." A learning and growth situation which places too great a demand upon any one of a child's readiness areas could result in the development of "blocks" in those other capacities where his maturity may have been equal to the tasks presented to him. Hedges reports teacher's observations that normal-aged children with mental abilities comparable to accelerated younger children seem to give less evidence of strain and frustration in their conduct at school than do the younger-aged pupils. It would seem, then, that the younger child is carrying a load that is disturbing to himself. Investigation by Hedges revealed that younger students seem to be bothered by restlessness, emotional upsets, fears, and tensions, lack of adequate neurophysiological coordination exhibited by a lack of interest in certain of the usual school activities.
In a study limited to boys, DiPasquale (1980) found that children who enter kindergarten at a younger age do not fare as well as older classmates. By confirming the birthdate effect, DiPasquale supported the view that some children are not ready for first grade work because of their ages. He emphasized the importance of early detection of at-risk children. Hamalainsen (1953) found that the underage child is at a disadvantage, particularly in social and emotional adjustment, which persisted through the grades. In a study of fourth graders, Dickinson and Larson (1963) found that younger children with higher intelligence quotients were less able to achieve than older children with lower intelligence quotients.

Weinstein (1968) observed that younger entrants have problems based on their difficulty in meeting the behavioral and academic expectations of their teachers and classmates. Any system which narrowed the gap between the young entrant's capabilities and what is expected of him should serve to lower the probability of persistent behavioral and academic problems. Weinstein stated further that younger first graders are less mature socially, physically, and emotionally. They were less ready for academic skills, less able to sit quietly, concentrate, follow directions, share teachers' attentions, as well as less able to meet their expectations. He
confirmed the hypothesis that emotionally-disturbed children are likely to have been the youngest entrants in their first grade classes.

Carroll (1963) compared third graders and concluded that older classmates had better persistence, attention span, self confidence, initiative, independence, social maturity, and acceptance. He found that even a few months of additional growth constituted an advantage when the child embarked upon formal education. The benefit a child derived from opportunities to acquire an understanding or skill depended to a large extent upon his level of maturity when the opportunities were provided. In learning something new, a child would gain as much competence from a shorter period of practice when he is older as he would gain from a longer period of exercise at an earlier age. Carroll postulated that the child who was under pressure to master learning tasks beyond his level of maturity faces compounded adjustment problems.

Kohlberg (1968) stated that while various forms of early stimulation and learning have value, they do not justify teaching things earlier that would come later with less effort. Specific types of preschool academic and linguistic training, even though immediately successful, were unlikely to have long-run, general, and beneficial effects. He saw the preschool need as one in which the
child had a qualitatively different mode of thought and orientation to the world than the older child, one which was a prelogical, preintellectual, or not ordered to external truth values.

In an intensive longitudinal study, King (1955) found that having attained a few additional months at the beginning of school was an important factor in a child's ability to meet imposed restrictions and tensions that school necessarily presented. Younger children had a continued disadvantage throughout elementary school. King wrote that younger entrants were more likely to show greater indications of poor personal and social adjustment to school. This study, as well as others, also confirmed that more boys lack readiness skills than do their female counterparts.

Among other authors who confirmed the disadvantages for children entering school early were Davis (1980), Maddux (1980), and Brown (1973). Davis, in a study for the state of Kentucky based on achievement tests given at the end of first, fourth, and eighth grades, found that six-year-old entrants did better than five-year-old entrants. Maddux carried this premise a step further by citing from his research the fact that a disproportionate number of early entrants are classified as learning disabled. Brown, in his research, concluded that although
some younger children were mentally advanced, they were held back by sheer lack of experiences. Additional time allowed experience which was a prime factor in perception and conceptualization.

In another study, Mewha (1976) found that older children generally do better. Wolff (1979), in an even stronger reaction, wrote that it is necessary to protect children from having to conform to a bell-shaped, biologically unsound curve, otherwise healthy, intelligent, but late-maturing children are at greater risk for school failure.

**Success Following an Additional Year of Kindergarten**

Many children have received the extra year needed for maturation. Gott (1964) found that younger groups achieved less than all older groups in all subjects at all grade levels. The study focused on the effect of a 9 to 11 month age difference at kindergarten entrance and achievement at the ends of second and sixth grades.

Halliwell's analysis (1964) indicated that early entrance to first grade resulted in lower achievement when compared with later entrants of similar abilities when given the same school experiences. Chronologically older children were significantly higher in academic achievement. They also had better eye coordination. This study centered
around the reading and nonreading activities in the fourth and fifth grade.

Drury (1981) maintained that an extra year increased readiness. Using a developmental placement program for children in school, she added, may enhance their school success.

Turley (1982) reported that many children, not retained in kindergarten, repeated a later grade. She found that those who spent a second year in kindergarten scored about eight months higher in reading and mathematics at the end of the first grade than those developmentally younger students who were promoted. Turley recommended that the practice of retaining children in kindergarten on the basis of carefully defined developmental immaturity was profitable for many children over a period of years.

In a similar study, Shatwell (1982) found the Gesell Developmental Test an effective instrument for identifying developmental maturity in children entering kindergarten. He also found that children retained in kindergarten on the basis of the Gesell findings were less likely to repeat subsequent grades. The immature child benefitted from two years in kindergarten as compared to the immature child who went on. If the nonready child did not repeat kindergarten, there was a high likelihood of retention in subsequent years. Recommending repetition of kindergarten
should be the initial and most important step in preventing failure for early entrants, wrote Donofrio (1977). Parents and teachers feeling the need to impinge on the child's basic needs, gained at his own pace, with his own behavioral age, must be realistically confronted rather than giving in to the caprices of an idealistic, environmentally oriented education system.

Finlayson (1975), after extensive testing and conferences with both parents and teachers, concluded that nonpromotion appeared to have no effect on the self-concept of first grade children. In fact, self-concept in the retainees increased, while those of immature students with borderline promotions suffered a decrease.

Scott and Ames (1969) also agreed that there was no emotional damage to an immature child who repeated first grade. They based these findings on results from a survey by parents of retained children who rated their children's social, emotional, physical, and academic adjustment for the year prior to repeating and the year the child repeated the grade. Chase (1968) further stated that the retained child liked school better. With more confidence and success, and better perceptual-motor abilities, the child developed to a point which approximated the experiences of the school system more closely than was the case in the year during which the failure occurred.
It is strongly suggested throughout the literature that readiness—the concept of the child's ability to engage in academic endeavors—is of prime importance in determining the child's academic success. The Gesell School Readiness Screening Test has been found to be a good predictor and a useful instrument for determining academic readiness. There is also evidence that certain of the Gesell subtests may be more useful than others in assisting those who must make placement decisions concerning young children.

**Gesell Subtests as Predictors**

Some researchers have discovered that certain subtests of the Gesell battery can be as accurate an indicator of readiness as the whole battery. Wasserman (1979) concluded that the Incomplete Man subtest was a good indicator of youngness. This subtest was also found to be culture free and could be used as a screening device if a modified scoring system was employed. Glennon (1978) found that teachers determined the Copy Forms and Incomplete Man subtests to be more useful than the other subtests.

Popovics (1982) studied the predictive validity of the five component actuarial scores of the Gesell Copy Forms subtest and concluded that the Proportion, Corners, and Closures scores provided reasonable long-term estimates of most of the selected achievement and intelligence
criterion measures and, as predictors, were better than, or generally as effective as, the Copy Forms total series. Criterion measures included the Iowa Test of Basic Skills, Lorge-Thorndike Intelligence Test, Goodnough-Harris Figure Drawing Test, and the Torrance Circle Test.

Heimgartner (1970), in a cross-sectional study of students at three grade levels, demonstrated the Copy Forms and Incomplete Man subtests of the Gesell battery yielded a correlation of .91 between developmental age and reading success. He concluded that these same subtests also gave the best prediction of developmental age and reading success. He further postulated that children should be tested yearly to find at what developmental age they are operating and subsequently grouped by developmental age. He found that the correlation of the ability to reproduce numbers and reading level was .90 and that reading level and the Incomplete Man subtest were correlated at .89.
CHAPTER 3
METHODS AND PROCEDURES

Procedures used in conducting the study and reviewed in this chapter include sample selection, instrument selection, data collection, and data analysis.

Population and Sample Selection
Initially this study was to involve a representative sample of 150 students from those who attended kindergarten in the Volusia County public schools during the 1982-83 school year. During that school year, the primary resource teachers were trained to administer the Gesell School Readiness Screening Test. Approximately 70 percent of the district's kindergarten students were tested with that instrument in the spring of 1983.

During the preplanning week in August, all primary resource teachers were surveyed to determine the number of kindergarten students tested in their schools. Four schools had their entire kindergarten population tested. With the assistance of the Director of Research and Accountability for the School Board of Volusia County, three of these schools were chosen for the project. The criteria used to ascertain that these schools were
representative of Volusia County Schools was the socio-economic status and majority-minority background. These three schools represented the Daytona Beach area. To make the study geographically representative of the district, two schools which had tested a higher percentage of their kindergarten students were added. One school was added from the southeastern section and one from the western section of Volusia County. The resulting sample of schools were representative of the three geographical areas of Volusia County. The students in these schools were representative of the entire student population of Volusia County public schools. This decision was based on the Director's personal knowledge of preparing the school district's demographic report from each school's survey.

The principals and primary resource teachers from the five centers agreed to their school's participation in this study. At each school 30 students were randomly selected from those kindergarten students tested during the 1982-83 school year.

During the spring of 1984, the primary resource teacher and principal of one of the elementary schools decided that the unique circumstances at that center prevented them from participating in the study. It was then determined that the school would have to be dropped from the study. With the loss of one center and the loss
of students who had moved from the other four centers, the sample was reduced to 72 students. According to the Director, these four schools were representative of the entire population of Volusia County. The extent to which the mortality rate destroyed the representativeness is unknown. However, it is assumed for the purposes of conducting the study, that it did not destroy the representativeness of the sample.

All four of the centers involved in the study included kindergarten classes, a kindergarten-first grade transitional class, and regular first grade classes in their early childhood program. The students who were judged to need an additional year to develop those readiness skills thought to be essential for achievement in first grade were placed in the transitional class. The transitional class was a planned environment of learning for those students who were six years old chronologically, but who were developmentally younger. These students were provided with an instructional mode that focused upon the sensory and experiential development.

Instruments

The three instruments that were used in the study were the Gesell School Readiness Screening Test, the Otis-Lennon School Ability Test, and the Comprehensive Test of Basic
Skills. All three tests were administered by the primary resource teacher at each school.

The Gesell School Readiness Screening Test is a test administered individually to young children to determine a developmental age. The test consisted of eight subtests: Cubes, Name, Copy Forms, Numbers, Incomplete Man, Initial Interview, Animals, and Interests. All of the Gesell subtests were administered to all participants from March 5 to June 5, 1983. Each subtest consisted of simple tasks and responses. The responses were matched with a set of norms derived by the Gesell Institute to determine the developmental age for each subtest. The developmental age was determined by subjective appraisal and reported in six-month intervals.

The Otis-Lennon School Abilities Test was used to determine the general intelligence test performance and was reported as a standard score (intelligence quotient). The scores were based on a normative mean of 100 and a standard deviation of 15 points. The IQ test was administered in January 1984. Although this test was chosen in September, it was not available for administration until January.

Reading achievement was measured with the reading section of Form U of the Comprehensive Test of Basic Skills. The test was administered during the first week in April 1984. Two levels of the test were used. Level B was
used for the students in the transitional-first grade class and Level C was administered to the students in the regular first grade class.

The achievement results were reported in **scale scores**. Scale scores are units of a single, equal-interval scale that is applied across all levels of Comprehensive Tests of Basic Skills Form U regardless of grade or time of year of testing. These scores are expressed as numbers that may range from 0 through 999. The equal-interval property of the scale makes scale scores especially appropriate for various statistical purposes. Scale scores can be added, subtracted, and averaged across test levels. Such computations permit direct comparisons among classes and schools (CTB/McGraw-Hill, 1982).

**Validity and Reliability**

According to Buros (1972), the Gesell School Readiness Screening Test manual presents the basic educational viewpoint of the Gesell Institute: that children should be entered in school and assigned on the basis of their developmental behavior age. In an initial study of 100 kindergarten children, the investigators' global judgments as to readiness were in close agreement with global teacher judgments, and also were highly related to academic section assignments six years later.
According to the Gesell test manual, responses are matched to a set of norms derived by examining thousands of children at every age level. The manual deals only with the Gesell Institute data and does not refer to other studies or relevance. The scoring procedure was very subjective and there is little literature that supports the reliability of the test. Because there is limited information on reliability, the validity of the instrument is in doubt.

Test reliability reflects the precision of the test as a measuring instrument, that is, the extent to which chance facts affect test results. The Otis-Lennon School Ability Test, Form R, used the Kuder-Richardson reliability procedure for the Primary 1 level. The Kuder-Richardson procedures yield an estimate of reliability that indicates the internal consistency of performance within the items of a test. For the norm group the Kuder-Richardson Formula 20 reliability coefficient was .90.

Test validity is generally defined as the extent to which the test measures what it is purported to measure. In order to determine validity, Otis-Lennon scores were correlated with teacher grades for elementary and high school curricular areas. The validity correlation for Primary 1 was .40 based on a sample size of 208 to 213. Additional evidence of validity could result from
determining the relationship between Otis-Lennon scores and achievement test scores, or data based on other accepted measures of general mental ability or scholastic aptitude.

The reliability and validity coefficients for the Comprehensive Test of Basic Skills were not available. However, according to the preliminary technical report, the validity will be based on content validity and the reliability measure will be calculated with the Kuder-Richardson procedure.

Data Collection

The primary resource teachers from the five schools involved in the study submitted their 1982-83 kindergarten lists to the author in September 1983. These lists consisted of students who had attended their respective schools and who were tested with the Gesell School Readiness Screening Test. A standard table of random numbers was used to determine the students to be included in the study.

This random-sample list of students was compiled and returned to each primary resource teacher. The following information was requested for each student: date of birth; sex; grade placement for the 1983-84 school year; date of Gesell Testing and the developmental age of each subtest in months; the IQ normed by age; and the test level and the scale score of each section of the Comprehensive Test of Basic Skills.
As tests became available from the publishing companies, they were forwarded to the primary resource teachers with a time line to insure uniform test administration. The Otis-Lennon School Ability Test was sent to each center in January 1984, and the Comprehensive Test of Basic Skills was sent in April 1984.

**Research Design and Treatment**

Stepwise multiple regression was used to examine the relationship among the subtests of the Gesell School Readiness Screening Test and sections of the reading achievement test.

The analysis included the following independent variables:

1. Subtests of the Gesell School Readiness Screening Test reported in six-month intervals by months; Cubes, Name, Copy Forms, Numbers, Incomplete Man, Initial Interview, Animals, Interests.

2. Sex reported as 0 for boys and 1 for girls.

3. Chronological age reported in months as of September 1, 1983.

4. Intelligence quotient reported by age norms.

5. Placement reported as 0 for transitional-first grade class and 1 for regular first grade class.
The dependent variables used in the research design were the scores on the reading sections of the Comprehensive Test of Basic Skills:

1. Word Attack
2. Vocabulary
3. Reading Comprehension
4. Language

The stepwise multiple regression program in Version X of the SPSS computer program was used to analyze the data. Frequency distributions and group means were also calculated.
The purpose of this study was to determine the degree to which the subtests of the Gesell School Readiness Screening Test administered at the kindergarten level predicted first grade reading achievement scores.

The 72 students involved in this study were comprised of 32 boys and 40 girls. The ages, as of September 1, 1983, ranged from 72 months to 83 months with a nearly equal distribution across age. Of the 72 students, 20 were placed in a transitional-first grade class and 52 in regular first grade classes. The IQ scores ranged from 67 to 136 with a mean of 98.5, a median IQ of 100, and a standard deviation of 17.6.

**Correlation**

In order to examine the relationships among the predictor variables, Pearson Product Moment Correlations were calculated among the dependent variables and between the predictor and dependent variables. Table 1 represents the upper triangular intercorrelation matrix among the independent or predictable variables. The variable sex had a dichotomous coding: 0 for boys and 1 for girls.
### TABLE 1

**INTERCORRELATIONS BETWEEN INDEPENDENT VARIABLES**

<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>Placement</th>
<th>IQ</th>
<th>Cubes</th>
<th>Copy Forms</th>
<th>Numbers</th>
<th>Incomplete Man</th>
<th>Interview</th>
<th>Animals</th>
<th>Interests</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>-.14</td>
<td>.16</td>
<td>.19</td>
<td>-.01</td>
<td>.00</td>
<td>.10</td>
<td>.20</td>
<td>.08</td>
<td>.07</td>
<td>.04</td>
<td>.14</td>
</tr>
<tr>
<td>Age</td>
<td>.20</td>
<td>-.19</td>
<td>.15</td>
<td>.10</td>
<td>.17</td>
<td>.09</td>
<td>.11</td>
<td>.31</td>
<td>.22</td>
<td>.02</td>
<td></td>
</tr>
<tr>
<td>Placement</td>
<td>.37**</td>
<td>.13</td>
<td>.22</td>
<td>.23</td>
<td>.25</td>
<td>.30</td>
<td>.22</td>
<td>.31</td>
<td>.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IQ</td>
<td>.22</td>
<td>.25</td>
<td>.39**</td>
<td>.35*</td>
<td>.40**</td>
<td>.33*</td>
<td>.17</td>
<td>.28</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cubes</td>
<td></td>
<td></td>
<td>.35*</td>
<td>.15</td>
<td>.20</td>
<td>.32</td>
<td>.23</td>
<td>.18</td>
<td>.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Copy Forms</td>
<td>.47**</td>
<td></td>
<td>.38**</td>
<td>.36*</td>
<td>.40**</td>
<td>.25</td>
<td>.54**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Numbers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.47**</td>
<td>.43**</td>
<td>.44**</td>
<td>.33*</td>
<td>.60**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incomplete Man</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.50**</td>
<td>.31</td>
<td>.28</td>
<td>.35*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interview</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.41**</td>
<td>.25</td>
<td>.45**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Animals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.37*</td>
<td>.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.28</td>
<td></td>
</tr>
</tbody>
</table>

**p < .01**

* p < .05
Age was coded in the actual number of months of age as of September 1, 1983. Placement was a dichotomous variable, coded 0 for the transitional class and 1 for the regular first grade class. The IQ was the actual intelligence quotient score obtained on the Otis-Lennon School Ability Test.

The correlations between these background variables were low, either low positive or low negative, all below .20 except the correlation between IQ and placement. Students placed in the regular first grade class tended to have a higher IQ than those students placed in the transitional-first grade class, although the correlation between placement and IQ was only moderate (.37).

The correlation found in Table 1 between the background variables and the Gesell subtests were all low with the exception of correlations involving IQ. None of the correlations between sex, age, and placement and any of the Gesell subtests exceeded .31. The IQ variable had a moderate correlation with Numbers, Incomplete Man, Interview, and Animals.

The intercorrelations among the eight Gesell subtests (also recorded in Table 1) tended to be in the moderate range, with many of them falling between .35 and .50. Only three correlations were at .50 or above, and these
were the correlations between Numbers and Name, Copy Forms and Name, and Incomplete Man and Interview. These correlations were important in interpreting the results of the stepwise multiple regression. Variables that had higher correlations tended to measure similar traits and the second variable was less likely to be included in the stepwise regression when a previous variable with which it was highly correlated was entered first.

A correlation matrix between all of the predictor variables and the four dependent measures--Word Attack, Vocabulary, Reading Comprehension, and Language--are found in Table 2. The third dependent measure, Reading Comprehension, had correlations based on 52 students rather than the full 72 because a Reading Comprehension score was produced only by test forms administered in the regular first grade.

As can be expected, the predictor variable that has the highest correlation with the dependent variables was IQ. This can be expected for two reasons: the closeness in time between the administration of the IQ and achievement test and the similarity between the concept of IQ and standardized achievement testing. Because only students placed in regular first grade had Reading Comprehension scores, the correlation between placement
## Table 2

Correlations between Predictor and Dependent Variables

<table>
<thead>
<tr>
<th></th>
<th>N=72 Word Attack</th>
<th>N=72 Vocabulary</th>
<th>N=52 Reading Comprehension</th>
<th>N=72 Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>.17</td>
<td>.21</td>
<td>.29</td>
<td>.12</td>
</tr>
<tr>
<td>Age</td>
<td>.09</td>
<td>.11</td>
<td>-.08</td>
<td>.03</td>
</tr>
<tr>
<td>Placement</td>
<td>.44</td>
<td>.63**</td>
<td>N/A</td>
<td>.58**</td>
</tr>
<tr>
<td>IQ</td>
<td>.68**</td>
<td>.62**</td>
<td>.47**</td>
<td>.55**</td>
</tr>
<tr>
<td>Cubes</td>
<td>.26</td>
<td>.23</td>
<td>.25</td>
<td>.25</td>
</tr>
<tr>
<td>Copy Forms</td>
<td>.19</td>
<td>.15</td>
<td>.16</td>
<td>.12</td>
</tr>
<tr>
<td>Numbers</td>
<td>.43**</td>
<td>.42**</td>
<td>.31</td>
<td>.41**</td>
</tr>
<tr>
<td>Incomplete Man</td>
<td>.36*</td>
<td>.25</td>
<td>.11</td>
<td>.27</td>
</tr>
<tr>
<td>Interview</td>
<td>.50**</td>
<td>.52**</td>
<td>.41**</td>
<td>.45**</td>
</tr>
<tr>
<td>Animals</td>
<td>.32</td>
<td>.40**</td>
<td>.41**</td>
<td>.45**</td>
</tr>
<tr>
<td>Interests</td>
<td>.23</td>
<td>.23</td>
<td>-.09</td>
<td>.20</td>
</tr>
<tr>
<td>Name</td>
<td>.34</td>
<td>.26</td>
<td>.15</td>
<td>.27</td>
</tr>
</tbody>
</table>

** * $p < .01$  
* $p < .05$
and Reading Comprehension was not applicable. The correlations between placement and the dependent variables were also moderately high as can be expected since a placement policy was in effect while the study was being conducted.

These correlations have two possible interpretations. One interpretation is that the students who were given regular first grade curriculum tended to score higher because they received the broader range of instruction. The second potential interpretation was that the placement procedure was moderately successful in placing students. Thus, students placed in the first grade were higher achievers than students placed in the interim or transitional first grade.

The correlations of student sex with each of the predictor variables were very low. For this reason, sex and age were not included as predictor variables in the stepwise linear regression.

Four of the Gesell subtests demonstrated relatively weak correlations with the predictor variables: Cubes, Copy Forms, Interests, and Name. The correlations indicated that scores on those four Gesell subtests had very little relationship to achievement test scores.

A fifth Gesell subtest, Incomplete Man, also demonstrated fairly weak correlations with the predictor
variables. Only one of these correlations, Incomplete Man and Word Attack ($r=0.36$), reached a level that could be considered a moderate correlation. Although the correlation was statistically significant at less than the 5 percent level, the magnitude of the correlation determines the strength of the relationship of the correlation.

The Gesell Interview subtest demonstrated the highest correlations with the predictor variables. No other Gesell subtest correlated higher than Interview with any of the four dependent measures.

The two remaining Gesell subtests, Numbers and Animals, had moderate correlations with each of the predictor variables. The relative size of the correlations for these two remaining predictors varied from variable to variable. But in no case did the correlations for either of these two variables exceed the demonstrated relationship of the Interview subtest with each of the predictor variables.

**Multiple Regression**

The answers for the basic questions for this study were obtained through a stepwise multiple regression. Two stepwise regression analysis procedures were used with each of the dependent variables. In the first regression procedure, all of the eight Gesell subtests were entered
into stepwise regression and the background variables, IQ and placement, were then entered in the second phase of the same regression analysis. This means that after all of the Gesell subtests were included in the regression equation, IQ and placement were added in an empirically based stepwise fashion. The stepwise multiple regression selects as its first predictor variable that one which has the highest zero order correlation with the dependent variable. Subsequent predictors were added based on the criterion of largest increment to the multiple correlation coefficient.

For the purposes of this study, we were interested in multiple correlation, the square of the multiple correlation at each step. The multiple correlation coefficient represented the correlation between each dependent variable and the composite set of variables included in the regression at each step. The square of the multiple correlation ($R^2$) represented the proportion of the variance in the dependent variable that was accounted for by the composite set of predictor variables at each step. The change in the $R^2$ represented the increase in the $R^2$ from step to step.

Because the multiple correlation values obtained from each study represented an overestimate of that likely to be obtained in the total population, each multiple regression description contained a report of an adjusted $R^2$ or multiple
correlation. The adjusted $R$ represented the expected correlation in the total population from which the sample is drawn. The adjusted correlation was always lower than the calculated correlation and was a conservative estimate.

The summary information is presented in Table 3 for the two stepwise multiple regression analyses conducted on the dependent variable Word Attack. The incremental $F$ test is reported in the table. The formula for the incremental $F$ test is:

$$F = \frac{(R_b^2 - R_a^2)/m}{(1 - R_b^2)/N - k - 1)}$$

$m =$ number of predictors added
$k =$ total number of predictors in $b$
$N =$ sample size
$R_b^2 =$ $R^2$ value at this step
$R_a^2 =$ $R^2$ value at previous step

As can be seen in the first of the two regression reports, only the Gesell subtests, Interview and Numbers, contributed a statistically significant amount to the prediction of the dependent score.

Because of all the Gesell subtests Interview correlated most highly with Word Attack, it was entered first. Numbers was entered second because it increased
### TABLE 3

**SUMMARY OF STEPWISE REGRESSION**

*FOR WORD ATTACK: GESELL SUBTESTS ENTERED FIRST*

<table>
<thead>
<tr>
<th>Predictor Variables</th>
<th>( R )</th>
<th>( R^2 )</th>
<th>( \Delta R^2 )</th>
<th>( \Delta R )</th>
<th>( \Delta R^2 )</th>
<th>( F )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Interview</td>
<td>0.505</td>
<td>0.255</td>
<td>0.255**</td>
<td>0.244</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2) Numbers</td>
<td>0.559</td>
<td>0.312</td>
<td>0.058*</td>
<td>0.293</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3) All Gesell</td>
<td>0.591</td>
<td>0.349</td>
<td>0.037</td>
<td>0.277</td>
<td></td>
<td>5.8</td>
</tr>
<tr>
<td>4) IQ</td>
<td>0.749</td>
<td>0.561</td>
<td>0.212**</td>
<td>0.505</td>
<td></td>
<td>29.86</td>
</tr>
<tr>
<td>5) Placement</td>
<td>0.768</td>
<td>0.589</td>
<td>0.029*</td>
<td>0.530</td>
<td></td>
<td>4.33</td>
</tr>
</tbody>
</table>

** Summary of stepwise regression for word attack: IQ and placement entered first **

<table>
<thead>
<tr>
<th>Predictor Variables</th>
<th>( R )</th>
<th>( R^2 )</th>
<th>( \Delta R^2 )</th>
<th>( \Delta R )</th>
<th>( \Delta R^2 )</th>
<th>( F )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) IQ</td>
<td>0.680</td>
<td>0.463</td>
<td>0.463**</td>
<td>0.455</td>
<td></td>
<td>60.13</td>
</tr>
<tr>
<td>2) Placement</td>
<td>0.711</td>
<td>0.505</td>
<td>0.042*</td>
<td>0.491</td>
<td></td>
<td>5.92</td>
</tr>
<tr>
<td>3) Interview</td>
<td>0.745</td>
<td>0.556</td>
<td>0.050**</td>
<td>0.536</td>
<td></td>
<td>7.68</td>
</tr>
<tr>
<td>4) All Gesell</td>
<td>0.768</td>
<td>0.589</td>
<td>0.033</td>
<td>0.544</td>
<td></td>
<td>0.703</td>
</tr>
</tbody>
</table>

** \( *P \leq 0.05 \)**

** \( **P \leq 0.01 \)**
the multiple correlation greater than any other predictor. None of the other Gesell subtests individually increased the prediction ability. After all Gesell subtests were entered, the variables IQ and placement were found to add a substantial increment to the prediction. Of the two variables, IQ added substantially more than placement.

As seen in the top half of Table 3, when the Gesell subtests were entered first, Interview had a $R^2$ value of .255 which means that about 25 percent of the variance in the dependent variable, Word Attack, was accounted for by Interview. Numbers increased the $R^2$ value by less than 6 percent so that the $R^2$ value was .312. When all of the additional Gesell subtests were added to the regression equation, the total $R^2$ increased by slightly less than .04. However, when IQ and placement were added, they added a statistically significant amount of approximately 24 percent (IQ added the larger amount). Even after all of the eight subtests were entered into the equation, the total $R^2$ value was slightly less than .35. The IQ score accounted for an additional 21 percent of the variance in Word Attack.

In order to determine if the Gesell subtests added information above and beyond what was available in IQ, the regression analysis was entered in another pattern. In this case, IQ and Placement were entered first. The results are
reported in the second half of Table 3. As seen in the table, IQ accounted for 46 percent of the variance in Word Attack scores and Placement accounted for 39 percent of the scores. When the single Gesell subtest: Interview, that was most highly correlated with residual was added to the equation, it accounted for slightly more than an additional 7 percent of the variance in Word Attack.

Table 4 provides two summary regression analyses for the dependent variable, Vocabulary. The findings followed a pattern similar to the dependent variable, Word Attack. The single exception was that after all the Gesell subtests were entered into the first analysis, the placement variable added more to the prediction of Vocabulary than did IQ. Thus, it was entered into the regression prior to IQ.

When the Gesell subtests were entered (top half of Table 4), Interview entered first and produced an $R^2$ value of .274; about 27 percent of the variance of Vocabulary was accounted for by Interview.

As in Word Attack, the predictor variable Numbers was added second and increased the $R^2$ value by less than 5 percent. The $R^2$ value increased to .323. None of the other Gesell subtests individually added a statistically significant amount to the prediction of Vocabulary. However, when placement and IQ were added, they added a
TABLE 4

SUMMARY OF STEPWISE REGRESSION
FOR VOCABULARY: GESELL SUBTESTS ENTERED FIRST

<table>
<thead>
<tr>
<th>Predictor Variables</th>
<th>$R$</th>
<th>$R^2$</th>
<th>$R^2$ Change</th>
<th>Adjusted $R^2$</th>
<th>Incremental $F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Interview</td>
<td>.523</td>
<td>.274</td>
<td>.274**</td>
<td>.263</td>
<td>$F_{(1,70)} = 26.35$</td>
</tr>
<tr>
<td>2) Numbers</td>
<td>.568</td>
<td>.323</td>
<td>.049*</td>
<td>.303</td>
<td>$F_{(1,69)} = 5.00$</td>
</tr>
<tr>
<td>3) All Gesell</td>
<td>.618</td>
<td>.381</td>
<td>.058</td>
<td>.314</td>
<td>$F_{(6,63)} = 0.99$</td>
</tr>
<tr>
<td>4) Placement</td>
<td>.784</td>
<td>.614</td>
<td>.233**</td>
<td>.565</td>
<td>$F_{(1,62)} = 36.98$</td>
</tr>
<tr>
<td>5) IQ</td>
<td>.832</td>
<td>.692</td>
<td>.078**</td>
<td>.647</td>
<td>$F_{(1,61)} = 15.60$</td>
</tr>
</tbody>
</table>

** $P < .01$
* $P < .05$

SUMMARY OF STEPWISE REGRESSION
FOR VOCABULARY: IQ AND PLACEMENT ENTERED FIRST

<table>
<thead>
<tr>
<th>Predictor Variables</th>
<th>$R$</th>
<th>$R^2$</th>
<th>$R^2$ Change</th>
<th>Adjusted $R^2$</th>
<th>Incremental $F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Placement</td>
<td>.627</td>
<td>.393</td>
<td>.393**</td>
<td>.385</td>
<td>$F_{(1,70)} = 45.39$</td>
</tr>
<tr>
<td>2) IQ</td>
<td>.754</td>
<td>.568</td>
<td>.175**</td>
<td>.555</td>
<td>$F_{(1,69)} = 27.88$</td>
</tr>
<tr>
<td>3) Interview</td>
<td>.789</td>
<td>.622</td>
<td>.054**</td>
<td>.605</td>
<td>$F_{(1,68)} = 9.705$</td>
</tr>
<tr>
<td>4) All Gesell</td>
<td>.834</td>
<td>.695</td>
<td>.073</td>
<td>.645</td>
<td>$F_{(7,61)} = 2.08$</td>
</tr>
</tbody>
</table>

** $P < .01$
* $P < .05$
statistically significant amount to the predictions. Placement added the larger amount, approximately 23 percent. A combined $R^2$ value of .381 was accounted for when all eight of the Gesell subtests were entered.

As in Word Attack, a second pattern was used in the regression analysis with Placement and IQ entered first. As illustrated in the second half of Table 4, Placement accounted for about 39 percent of the variance in the Vocabulary scores by itself. IQ accounted for an additional 18 percent of the variance in the scores. When the Gesell subtest Interview was added to the equation, it accounted for slightly more than 5 percent of the variance in Vocabulary. Placement, IQ, and Interview were all statistically significant at the 1 percent level.

The regression analysis for Reading Comprehension scores followed a similar pattern. It should be noted, however, that the regression on Reading Comprehension was conducted only on the data from the 52 students placed in the regular first grade. The data in Table 5 summarize the results of the two regression analyses. As demonstrated in the table, only one Gesell subtest, Interview, contributed a statistically significant amount to the prediction of Reading Comprehension scores.

When all of the Gesell subtests were entered first, Interview had a $R^2$ value of .257. Therefore, about 26
TABLE 5
SUMMARY OF STEPWISE REGRESSION
FOR READING COMPREHENSION: GESELL SUBTESTS ENTERED FIRST

<table>
<thead>
<tr>
<th>Predictor Variables</th>
<th>$R$</th>
<th>$R^2$</th>
<th>$R^2$ Change</th>
<th>Adjusted $R$</th>
<th>Incremental $F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Interview</td>
<td>.507</td>
<td>.257</td>
<td>.257**</td>
<td>.242</td>
<td>$F(1,50) = 17.25$</td>
</tr>
<tr>
<td>2) All Gesell</td>
<td>.645</td>
<td>.416</td>
<td>.159</td>
<td>.304</td>
<td>$F(7,43) = 1.67$</td>
</tr>
<tr>
<td>3) IQ</td>
<td>.712</td>
<td>.507</td>
<td>.092**</td>
<td>.399</td>
<td>$F(1,42) = 7.86$</td>
</tr>
</tbody>
</table>

** $P < .01$
* $P < .05$

SUMMARY OF STEPWISE REGRESSION
FOR READING COMPREHENSION: IQ ENTERED FIRST

<table>
<thead>
<tr>
<th>Predictor Variables</th>
<th>$R$</th>
<th>$R^2$</th>
<th>$R^2$ Change</th>
<th>Adjusted $R$</th>
<th>Incremental $F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) IQ</td>
<td>.477</td>
<td>.227</td>
<td>.227**</td>
<td>.211</td>
<td>$F(1,50) = 14.65$</td>
</tr>
<tr>
<td>2) Interview</td>
<td>.607</td>
<td>.369</td>
<td>.142**</td>
<td>.342</td>
<td>$F(1,49) = 11.01$</td>
</tr>
<tr>
<td>3) All Gesell</td>
<td>.712</td>
<td>.507</td>
<td>.138</td>
<td>.399</td>
<td>$F(7,42) = 1.68$</td>
</tr>
</tbody>
</table>

** $P < .01$
* $P < .05$
percent of the variance in the dependent variable, Reading Comprehension, was accounted for by Interview. When the other seven subtests were added to the regression equation, the total $R^2$ above that which was accounted for by Interview was .159. When IQ was entered, it was statistically significant, accounting for an additional 9 percent of the variance in Reading Comprehension scores.

A similar pattern was found in the second half of Table 5 when IQ was entered into the regression equation first. The predictor variable IQ accounted for about 23 percent of the variance in the Reading Comprehension scores by itself. Interview added a statistically significant amount to $R^2$, about 14 percent. In combination the seven remaining subtests accounted for only about 14 percent to the variance, a collective increase that was not significant.

The summaries of the two stepwise regression analyses for the dependent variable, Language, are included in Table 6. The results illustrated in this table show a pattern similar to that for Word Attack. When all the Gesell subtests were entered first, only the subtests Interview and Numbers contributed a statistically significant amount to the prediction of Language scores.

Interview produced an $R^2$ value of .20 and Numbers added an additional 6 percent so that the $R^2$ value for both
### TABLE 6
SUMMARY OF STEPWISE REGRESSION
FOR LANGUAGE: GESELL SUBTESTS ENTERED FIRST

<table>
<thead>
<tr>
<th>Predictor Variables</th>
<th>R</th>
<th>$R^2$</th>
<th>$R^2$ Change</th>
<th>Adjusted $R^2$</th>
<th>Incremental $F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Interview</td>
<td>.448</td>
<td>.200</td>
<td>.200**</td>
<td>.189</td>
<td>$F(1,70) = 17.54$</td>
</tr>
<tr>
<td>2) Numbers</td>
<td>.510</td>
<td>.261</td>
<td>.060*</td>
<td>.239</td>
<td>$F(1,69) = 5.61$</td>
</tr>
<tr>
<td>3) All Gesell</td>
<td>.558</td>
<td>.311</td>
<td>.050</td>
<td>.259</td>
<td>$F(6,63) = 0.76$</td>
</tr>
<tr>
<td>4) Placement</td>
<td>.720</td>
<td>.518</td>
<td>.207**</td>
<td>.474</td>
<td>$F(1,62) = 26.54$</td>
</tr>
<tr>
<td>5) IQ</td>
<td>.754</td>
<td>.569</td>
<td>.051**</td>
<td>.522</td>
<td>$F(1,61) = 7.18$</td>
</tr>
</tbody>
</table>

** $P < .01$  
* $P < .05$

### SUMMARY OF STEPWISE REGRESSION
FOR LANGUAGE: IQ AND PLACEMENT ENTERED FIRST

<table>
<thead>
<tr>
<th>Predictor Variables</th>
<th>R</th>
<th>$R^2$</th>
<th>$R^2$ Change</th>
<th>Adjusted $R^2$</th>
<th>Incremental $F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Placement</td>
<td>.583</td>
<td>.340</td>
<td>.340**</td>
<td>.331</td>
<td>$F(1,70) = 36.07$</td>
</tr>
<tr>
<td>2) IQ</td>
<td>.686</td>
<td>.471</td>
<td>.130**</td>
<td>.455</td>
<td>$F(1,69) = 16.99$</td>
</tr>
<tr>
<td>3) Interview</td>
<td>.710</td>
<td>.504</td>
<td>.033*</td>
<td>.482</td>
<td>$F(1,68) = 4.56$</td>
</tr>
<tr>
<td>4) All Gesell</td>
<td>.760</td>
<td>.578</td>
<td>.074*</td>
<td>.517</td>
<td>$F(7,61) = 2.17$</td>
</tr>
</tbody>
</table>

** $P < .01$  
* $P < .05$
subtests was .261. None of the other Gesell subtests individually added a statistically significant amount to the prediction of Language. However, when placement and IQ were added, they individually added a statistically significant amount. Placement added the larger amount, approximately 21 percent.

As seen in the summary of the second regression analysis in Table 6, after IQ was entered into the regression equation only Interview added a statistically significant amount to the prediction of the variable Language. Placement accounted for 34 percent and IQ accounted for about 13 percent of the variance in the Language scores. Interview was added and accounted for slightly more than 3 percent of the variance in the Language scores. It can be noted, however, when all of the remaining Gesell subtests were added, the increase in $R^2$ was statistically significant at the 5 percent level ($F_{(7,61)} = 2.17$). The calculated $F$ value was exactly equal to the critical ratio contained in the $F$ table. However, as a group the remaining Gesell subtests added only a slight increment to the prediction value. None of the seven subtests were independently statistically significant.

Descriptive Review

In order to provide more detailed information for the purposes of interpretation and recommendations to the
school district, several additional tables and graphs were produced. A set of eight tables (Tables 7 through 14) provided a mean and standard deviation on the four dependent measures for the students who scored at each level of a Gesell subtest.

When examining the subtest Cubes (Table 7), there was a substantial break in the mean score on each of the four subtests for those students who scored 60 or above and those students who scored 54. There was also a large break in the mean score between those students who scored 54 and those who scored below 54. A similar break was evident in the mean scores of the four dependent measures for individuals scoring at each level of the subtest Copy Forms (Table 9).

There was not a consistent or marked difference in the mean scores of the dependent variables for students who scored at the two levels, 60 or 66 and above, for either of the two Gesell subtests, Copy Forms and Cubes. The most consistent level of discrimination of achievement for subtests Copy Forms and Cubes was between 54 and 60, rather than 60 and 66.

On the Interest subtest (Table 14), no student scored below 60. There was a noteworthy discrimination between the means of the students who scored 60 and those who scored above 60 for all of the dependent measures except
<table>
<thead>
<tr>
<th>Cubes Score</th>
<th>n</th>
<th>Word Attack</th>
<th>Vocabulary</th>
<th>Language</th>
<th>Reading Comprehension</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(x) (SD)</td>
<td>(x) (SD)</td>
<td>(x) (SD)</td>
<td>n</td>
</tr>
<tr>
<td>72</td>
<td>12</td>
<td>464 (100)</td>
<td>436 (71)</td>
<td>409 (101)</td>
<td>9</td>
</tr>
<tr>
<td>66</td>
<td>20</td>
<td>472 (127)</td>
<td>459 (97)</td>
<td>425 (136)</td>
<td>15</td>
</tr>
<tr>
<td>60</td>
<td>19</td>
<td>456 (126)</td>
<td>453 (87)</td>
<td>413 (129)</td>
<td>14</td>
</tr>
<tr>
<td>54</td>
<td>16</td>
<td>416 (130)</td>
<td>399 (116)</td>
<td>366 (129)</td>
<td>12</td>
</tr>
<tr>
<td>48</td>
<td>4</td>
<td>333 (49)</td>
<td>367 (56)</td>
<td>297 (105)</td>
<td>2</td>
</tr>
<tr>
<td>42</td>
<td>1</td>
<td>320 ( - )</td>
<td>345 ( - )</td>
<td>178 ( - )</td>
<td>1</td>
</tr>
</tbody>
</table>
TABLE 8

COMPARISON OF ACHIEVEMENT TEST SECTION
FOR EACH DEVELOPMENTAL AGE OF THE GESELL NAME TEST

<table>
<thead>
<tr>
<th>Name</th>
<th>Word Attack</th>
<th>Vocabulary</th>
<th>Language</th>
<th>Reading Comprehension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score</td>
<td>n (x) (SD)</td>
<td>(x) (SD)</td>
<td>(x) (SD)</td>
<td>n (x) (SD)</td>
</tr>
<tr>
<td>72</td>
<td>11 473 (117)</td>
<td>419 (130)</td>
<td>417 (111)</td>
<td>8 415 (115)</td>
</tr>
<tr>
<td>66</td>
<td>39 475 (115)</td>
<td>470 (75)</td>
<td>428 (124)</td>
<td>34 420 (137)</td>
</tr>
<tr>
<td>60</td>
<td>20 377 (114)</td>
<td>377 (83)</td>
<td>324 (127)</td>
<td>8 429 (108)</td>
</tr>
<tr>
<td>54</td>
<td>1 428 (-)</td>
<td>430 (-)</td>
<td>345 (-)</td>
<td>1 418 (-)</td>
</tr>
<tr>
<td>48</td>
<td>1 250 (-)</td>
<td>315 (-)</td>
<td>366 (-)</td>
<td>1 120 (-)</td>
</tr>
<tr>
<td>Copy Forms Score</td>
<td>Word Attack $n$</td>
<td>(x) (SD)</td>
<td>Vocabulary $\bar{x}$ (SD)</td>
<td>Language $\bar{x}$ (SD) $n$</td>
</tr>
<tr>
<td>------------------</td>
<td>----------------</td>
<td>----------------</td>
<td>----------------</td>
<td>----------------</td>
</tr>
<tr>
<td>72</td>
<td>13</td>
<td>434 (113)</td>
<td>415 (117)</td>
<td>376 (86)</td>
</tr>
<tr>
<td>66</td>
<td>32</td>
<td>471 (127)</td>
<td>460 (84)</td>
<td>423 (141)</td>
</tr>
<tr>
<td>60</td>
<td>22</td>
<td>439 (119)</td>
<td>426 (95)</td>
<td>398 (122)</td>
</tr>
<tr>
<td>54</td>
<td>4</td>
<td>334 (56)</td>
<td>353 (34)</td>
<td>237 (82)</td>
</tr>
<tr>
<td>48</td>
<td>1</td>
<td>250 (-)</td>
<td>315 (-)</td>
<td>366 (-)</td>
</tr>
<tr>
<td>Numbers Score</td>
<td>Word Attack</td>
<td>Vocabulary</td>
<td>Language</td>
<td>Reading Comprehension</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------</td>
<td>------------</td>
<td>----------</td>
<td>-----------------------</td>
</tr>
<tr>
<td></td>
<td>n (x) (SD)</td>
<td>(x) (SD)</td>
<td>(x) (SD)</td>
<td>n (x) (SD)</td>
</tr>
<tr>
<td>72</td>
<td>8 570 (76)</td>
<td>507 (47)</td>
<td>493 (83)</td>
<td>8 461 (65)</td>
</tr>
<tr>
<td>66</td>
<td>43 450 (111)</td>
<td>449 (92)</td>
<td>413 (124)</td>
<td>33 419 (137)</td>
</tr>
<tr>
<td>60</td>
<td>15 401 (125)</td>
<td>380 (97)</td>
<td>341 (126)</td>
<td>7 422 (134)</td>
</tr>
<tr>
<td>54</td>
<td>3 346 (153)</td>
<td>364 (79)</td>
<td>261 (88)</td>
<td>2 333 (121)</td>
</tr>
<tr>
<td>48</td>
<td>3 324 (84)</td>
<td>354 (44)</td>
<td>293 (108)</td>
<td>2 215 (134)</td>
</tr>
<tr>
<td>Incomplete Man</td>
<td>Word Attack</td>
<td>Vocabulary</td>
<td>Language</td>
<td>Reading Comprehension</td>
</tr>
<tr>
<td>----------------</td>
<td>-------------</td>
<td>------------</td>
<td>----------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>Score</td>
<td>n</td>
<td>(x) (SD)</td>
<td>(x) (SD)</td>
<td>(x) (SD)</td>
</tr>
<tr>
<td>72</td>
<td>6</td>
<td>454 (121)</td>
<td>461 (65)</td>
<td>392 (83)</td>
</tr>
<tr>
<td>66</td>
<td>33</td>
<td>488 (107)</td>
<td>450 (101)</td>
<td>426 (129)</td>
</tr>
<tr>
<td>60</td>
<td>29</td>
<td>416 (119)</td>
<td>424 (90)</td>
<td>389 (120)</td>
</tr>
<tr>
<td>54</td>
<td>4</td>
<td>269 (74)</td>
<td>336 (80)</td>
<td>194 (51)</td>
</tr>
<tr>
<td>Interview Score</td>
<td>Word Attack</td>
<td>Vocabulary</td>
<td>Language</td>
<td>Reading Comprehension</td>
</tr>
<tr>
<td>-----------------</td>
<td>-------------</td>
<td>------------</td>
<td>----------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>78</td>
<td>568 (–)</td>
<td>598 (–)</td>
<td>606 (–)</td>
<td>591 (–)</td>
</tr>
<tr>
<td>72</td>
<td>543 (77)</td>
<td>500 (26)</td>
<td>472 (72)</td>
<td>461 (60)</td>
</tr>
<tr>
<td>66</td>
<td>478 (110)</td>
<td>462 (81)</td>
<td>431 (125)</td>
<td>438 (115)</td>
</tr>
<tr>
<td>60</td>
<td>366 (110)</td>
<td>369 (95)</td>
<td>310 (106)</td>
<td>358 (136)</td>
</tr>
<tr>
<td>54</td>
<td>316 (80)</td>
<td>333 (25)</td>
<td>305 (57)</td>
<td>120 (–)</td>
</tr>
<tr>
<td>48</td>
<td>359 (–)</td>
<td>387 (–)</td>
<td>383 (–)</td>
<td>120 (–)</td>
</tr>
</tbody>
</table>
### TABLE 13

**COMPARISON OF ACHIEVEMENT TEST SECTION**

**FOR EACH DEVELOPMENTAL AGE OF THE GESELL ANIMALS TEST**

<table>
<thead>
<tr>
<th>Animals Score</th>
<th>Word Attack</th>
<th>Vocabulary</th>
<th>Language</th>
<th>Reading Comprehension</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\bar{x}) (SD)</td>
<td>(\bar{x}) (SD)</td>
<td>(\bar{x}) (SD)</td>
<td>(\bar{x}) (SD)</td>
</tr>
<tr>
<td>84</td>
<td>649 ( - )</td>
<td>515 ( - )</td>
<td>479 ( - )</td>
<td>493 ( - )</td>
</tr>
<tr>
<td>72</td>
<td>476 (116)</td>
<td>475 (80)</td>
<td>447 (124)</td>
<td>446 (101)</td>
</tr>
<tr>
<td>66</td>
<td>466 (122)</td>
<td>459 (78)</td>
<td>409 (127)</td>
<td>429 (141)</td>
</tr>
<tr>
<td>60</td>
<td>390 (114)</td>
<td>375 (102)</td>
<td>346 (124)</td>
<td>367 (143)</td>
</tr>
<tr>
<td>54</td>
<td>440 ( - )</td>
<td>387 ( - )</td>
<td>270 ( - )</td>
<td>247 ( - )</td>
</tr>
<tr>
<td>42</td>
<td>375 ( - )</td>
<td>387 ( - )</td>
<td>366 ( - )</td>
<td>282 ( - )</td>
</tr>
</tbody>
</table>
TABLE 14

COMPARISON OF ACHIEVEMENT TEST SECTION
FOR EACH DEVELOPMENTAL AGE OF THE GESELL INTERESTS TEST

<table>
<thead>
<tr>
<th>Interests</th>
<th>Score</th>
<th>Word Attack</th>
<th>Vocabulary</th>
<th>Language</th>
<th>Reading Comprehension</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>((\bar{x})) (SD)</td>
<td>((\bar{x})) (SD)</td>
<td>((\bar{x})) (SD)</td>
<td>n</td>
</tr>
<tr>
<td></td>
<td></td>
<td>((\bar{x})) (SD)</td>
<td>((\bar{x})) (SD)</td>
<td>((\bar{x})) (SD)</td>
<td>n</td>
</tr>
<tr>
<td></td>
<td></td>
<td>((\bar{x})) (SD)</td>
<td>((\bar{x})) (SD)</td>
<td>((\bar{x})) (SD)</td>
<td>n</td>
</tr>
<tr>
<td></td>
<td></td>
<td>((\bar{x})) (SD)</td>
<td>((\bar{x})) (SD)</td>
<td>((\bar{x})) (SD)</td>
<td>n</td>
</tr>
<tr>
<td></td>
<td></td>
<td>((\bar{x})) (SD)</td>
<td>((\bar{x})) (SD)</td>
<td>((\bar{x})) (SD)</td>
<td>n</td>
</tr>
<tr>
<td>78</td>
<td>1</td>
<td>649 (- )</td>
<td>515 (- )</td>
<td>479 (- )</td>
<td>1</td>
</tr>
<tr>
<td>72</td>
<td>3</td>
<td>448 (183)</td>
<td>426 ( 80)</td>
<td>405 (174)</td>
<td>2</td>
</tr>
<tr>
<td>66</td>
<td>40</td>
<td>459 (110)</td>
<td>453 ( 92)</td>
<td>416 (121)</td>
<td>34</td>
</tr>
<tr>
<td>60</td>
<td>28</td>
<td>414 (129)</td>
<td>404 ( 98)</td>
<td>362 (133)</td>
<td>15</td>
</tr>
</tbody>
</table>
Reading Comprehension. The students who scored at 60 in Interests had a higher mean than students who scored at 66 and 72.

For the Gesell subtest Name (Table 8), only two students scored below 60. Therefore, the means for those two students cannot be considered for generalizability. However, there was a notable difference in the mean scores for those who scored 66 and above and those who scored 60 for each of the dependent variables, except Reading Comprehension.

One post factum interpretation might be that the placement decisions were made at the break between 60 and 66. Placement would reflect in all of the dependent variables, except Reading Comprehension. If this interpretation is valid, the differential in achievement scores may have represented placement decisions more than other factors.

As noted in Table 12, the Gesell subtest Interview had a wider range of scores than most of the other predictors. This wider range may in part account for its higher correlation with the dependent variables. For each of the dependent variables, the mean score for students scoring 66 or higher was notably greater than those who scored 60 or below.
The pattern of the means of the dependent variables for the subtest Animals (Table 13) followed a pattern fairly similar to the other predictor variables with the exception of the dependent variable Language. There were no noteworthy breaks between dependent variable means between those scoring 66 and above and those scoring 60 and below. For the Language dependent variable, there were noteworthy breaks between 72 and 66 and between scores of 66 and 60.

The tables of means for the Gesell subtest Numbers (Table 10) did not present a clear pattern of breaks by developmental score. The pattern varied by dependent measure.

The means of the dependent measures for each level obtained by groups of students on the subtest Incomplete Man (Table 11) also presented a complex picture. There were breaks of different magnitude on each of the four dependent variables. Those who had the highest scores on Incomplete Man, that is a score of 72, tended to score lower on the dependent variables, Word Attack, Reading Comprehension, and Language, than those who scored 66. However, those scores were based on only six students.

It has been the custom for those students scoring 66 or higher to be placed in the regular first grade. To help review the pattern of achievement of each dependent
variable for students who scored above and below this criterion, a series of graphs were developed.

These figures were included in the Appendix. As observed when viewing the figures, there was a very similar distribution of achievement test scores for students scoring above and below the cutoff. The figures helped describe the reason for poor, moderate, and low correlations between the Gesell subtests and the dependent variables.
CHAPTER 5
SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

The Volusia County School District has been involved in providing the appropriate learning environment for children in the early years through the use of a transitional kindergarten-first grade class. Students were placed in this class who were chronologically aged six but who were developmentally younger. The Gesell School Readiness Screening Test was used in making placement decisions.

During the 1982-83 school year, the primary resource teachers in the district were trained in the use of the Gesell School Readiness Screening Test. The Gesell test consists of eight subtests and is individually administered. The testing time usually consists of 45 minutes of working with the student and an additional 45 minutes for interpretation.

This study would assist the district in two ways. If certain subtests of the Gesell test are found to be better predictors of reading achievement, more students could be tested using fewer subtests. Also, there would be more accurate decisionmaking ability in regard to placement when there is wide disparity in the subtest scores.
Summary and Conclusions

The major purpose of this study was to determine which subtest of the Gesell School Readiness Screening Test is the best predictor of reading achievement. The procedure used was a stepwise regression to determine which subtest was the best predictor of achievement for each reading section of the Comprehensive Test of Basic Skills.

Seventy-two students from four schools were evaluated for this study. Three testing instruments were administered to these students. They were the Gesell test, the Otis-Lennon test, and the Comprehensive Test of Basic Skills.

As indicated in the review of literature, developmental level is widely accepted as an indicator of school readiness (Brenner and Scott, 1971; Ogletree, 1973; Carll and Richard, 1977; Wood, 1981; Weiss, 1962; Heffernan, 1962). Several researchers have found that scores on the Gesell and other indicators of readiness to be related to reading success and other school performance when the measures are taken concurrently (Andrews, 1971; Wood, 1981; Shatwell, 1982).

There may well be, and perhaps is, a strong relation between developmental levels and a child's concurrent ability to benefit from specific forms of instruction. School programs at all levels, especially the kindergarten
and first grade levels, need to pay particular attention to matching instruction to student needs and readiness. Yet there is little evidence that indicators of readiness such as the Gesell subtests serve as adequate long-term predictors of student performance. As demonstrated by the data of this study, student academic achievement levels one year after Gesell testing in the spring of their kindergarten year is only moderately related to early developmental levels.

Heimgartner and other researchers studying developmental level and school achievement with concurrent measures found the highest correlations between the Incomplete Man subtest of the Gesell and reading scores. The concurrent study found the Incomplete Man subtest of the Gesell to have a very low correlation with achievement scores one year later. Of all of the eight Gesell subtests: Cubes, Name, Copy Forms, Numbers, Incomplete Man, Interview, Animals, and Interests, the best single predictor of school achievement in this study was Interview.

An important consideration is that the Interview subtest had some predictive ability for reading achievement. The subtest Interview was a short verbal exercise consisting of questions usually within the realm of the child's experience: his age, birth date, last birthday, brothers and sisters, and parents' occupations.
The response from the child gave the examiner a quick glimpse of the child's level of performance, powers of organization, tempo, ability to stay on task, and the level of verbal language development. The child was asked to specify as much as he could without being pressured. When at an impasse, the examiner shifted the questioning or proceeded to the next item. The examiner recorded what the child said, exactly as it was stated. The child's sentence construction, grammatical errors, and pronunciation told a story about the development of the child.

Data for Heimgartner's study were collected for 104 students attending the University of Northern Colorado Laboratory School, Greeley, Colorado. The subjects were kindergarten to third grade students who were above average in mean intelligence with a mean IQ on the individual WISC of 114. The test for the study was the School Readiness Behavior Tests used at the Gesell Institute. The reading level of the child at the time of the developmental testing was selected as a criterion. The predictor variables used were sex, chronological age, composite developmental age, and intelligence quotient.

Heimgartner tested the hypothesis that developmental age is a good predictor of reading success. His regression analysis included all of the predictor variables. He found that when all of the variables (sex, chronological age,
composite developmental age, and IQ) were used, an $R^2$ of .84 was obtained. To determine the significance of developmental age as a predictor of reading success, a restricted model was used eliminating developmental age. With the restricted model the $R^2$ value computed was .51. The drop in $R^2$ from the full model to the restricted model was .30, which gave an $F$ value of 199.97. The $F$ value was significant at the one-tenth of one percent level of significance.

A secondary purpose of Heimgartner's study was to determine which Gesell subtests would be the best predictors of reading success. Using the individual student's reading level at the time of testing as the criterion for reading success and the Gesell subtests as predictors, he found that an $R^2$ of .87 was obtained when Incomplete Man was used as a single predictor. When Incomplete Man and Copy Forms were used together, an $R^2$ of .91 was obtained. When the Gesell subtests were used as predictors, the $R^2$ was higher than when the Gesell composite was used with other predictors. His final determination was that these two subtests were the best indicators of developmental age and reading success.

Heimgartner used the term reading success for the reading level of the student at the time of the developmental testing. He gave a rather specialized
meaning to the term prediction. He did not conduct a prediction study in the sense of predicting a future score. Heimgartner's study can be classified as concurrent validity investigation. He identified the Gesell variables that correlate with reading level when both tests are given at approximately the same time. He did not examine prediction as school districts normally use the term.

Other limitations of Heimgartner's study were the use of students from four grade levels and a lack of information presented on IQ testing. He reported the use of a sample with a mean IQ of 114 and a standard deviation of 17.6. However, information on when the IQ test was administered, by whom, and under what conditions was not reported. The use of students from four grade levels produced a high prediction that would not have been obtained for students at only one grade level.

The present study examined the extent to which Gesell subtests predicted reading achievement one year later. Interview was the strongest predictor of scores on each of four sections of the achievement test. When the Gesell subtests were entered into the regression (Table 11 through 14), Interview produced an $R^2$ of .26 with Word Attack, .27 with Vocabulary, .26 with Reading Comprehension, and .20 with Language. After the information for Interview was entered into the regression equation, the remaining seven
subtest scores added very little to the prediction. Because the Incomplete Man subtest was moderately correlated with Interview (Table 1), any predictive ability that Incomplete Man may have had was accounted for by the early entry of Interview. The correlations between Incomplete Man and the dependent measures (Table 2) were weak.

The frequency polygons for spring, 1984 achievement scores visually demonstrated the weak predictability of Gesell subtests. When the spring, 1984 achievement score distribution for students scoring below the criterion of 66 months in the spring of 1983 are compared with those of students scoring above the criterion, there is little difference (see Appendix). The patterns were developed to illustrate the achievement of those students who scored above 66 months and those who scored 66 months and below. Usually students who scored 66 and above were placed in the regular first grade class. There was a tremendous amount of overlap in the achievement performance in those who passed and those who did not pass the cutoff point for placement in the transitional and regular first grade class.

There are many potential post-factum interpretations of these results. One explanation of the inability to obtain high prediction is the low amount of demonstrated reliability and validity of Gesell scores. Another is
differential levels and rates of development for individual children. Children may experience strong but irregular developmental spurts. Tests taken during the spring of the kindergarten year may have immediately followed a developmental spurt for some children and immediately preceded developmental spurts for others. Other children may show substantial developmental growth during the summer months and early parts of the following fall.

The data for this study indicate that placement of students in the transitional first grade class or regular grade was a moderately strong predictor of student achievement scores. This finding is not unusual because placement was based in part on developmental level of the child at the end of kindergarten and also because of a differential instructional program. The instructional program probably influenced much of the relationship between placement and student achievement because placement contributed a noteworthy increase to the $R^2$ after all of the Gesell subtests had been entered into a regression. Thus placement had a significant partial correlation with achievement test scores after the Gesell subtests were partialed from the relationship.

There are indications that the correlations between predictor variables and the dependent measures were affected in part by the six-month interval scores. Because the
scoring methodology is subjective, it can be expected to have low reliability. There is little research information that suggests the scoring of the Gesell test has accepted reliability.

Although there was some predictive ability in the Gesell School Readiness Screening Test, the predictive ability was not strong. IQ was more strongly related to achievement. Three factors may account for the low predictability by the Gesell subtests and higher predictability by IQ. First, the primary resource teachers were newly certified and may not have been comfortable with the test. In fact, some of these tests were administered before the administrators were fully certified. Second, the correlation between the IQ and achievement test was relatively high because the IQ test was given relatively close in time to the achievement test. The IQ test was given in January, almost an average of nine months after the Gesell test. Since IQ essentially measures the same factor as an achievement test, the IQ test was equivalent to a general academic achievement test. Therefore, this broad-range academic test was given at a time close to that of the narrower achievement test. If the IQ test had been given approximately the same time as the Gesell test, the correlation may have been lower. Third, the fact that the Gesell subtests were scored in six-month intervals rather
than in smaller increments did not allow for high
discrimination between developmental levels. There was a
coarse level of discrimination between levels of student
development and consequently the Gesell test had a low
potential for a high correlation.

The results of this study do not indicate that the
Gesell instrument should be discarded as a screening device
for determining immediate action on student instruction.
However, when such information as the Gesell indicators are
used in placing children in instructional programs, the
programs need to be designed in such a way that teachers
can make adjustment for students with different
developmental growth patterns. Although Interview
demonstrated the best prediction with school achievement in
this study, other studies using Gesell indicators as
concurrent readiness indicate that Incomplete Man and Copy
Forms are better indicators of students' readiness for
various reading levels. Volusia County might want to
review the current strategy of using Gesell testing as an
indicator for placement for children who are leaving the
kindergarten levels. However, the Gesell may still be very
useful as an indicator of readiness at various times during
the year. One problem with using the Gesell test in the
latter manner is its length of administration and scoring
time. However, other studies may indicate that specific
subtests of the Gesell test such as Incomplete Man and Copy Forms may in fact be able to substitute for the complete Gesell test for concurrent scoring for instructional planning and student placement purposes.

Recommendations

Based on the research, observations, conclusions, and limitations of the study, several recommendations are submitted.

If the district plans to continue to implement a transitional-first grade class, given the limitations of these data, the following recommendations should be considered.

1. The district should not put a great deal of emphasis on the results of the Gesell test. Gesell scores are now being used as a primary influence on grade placement.

2. If only one subtest of the Gesell test is to be administered, Interview should be used.

3. The district should conduct an intensive study to determine if the transitional-first grade class is beneficial for students. Such a study should identify the characteristics of students that may benefit from placement in the transitional-first grade.

This study had several limitations. The findings need to be verified by additional research. Such research
should consider the following recommendations:

1. Use a larger number of students to determine whether or not the same results would be found.

2. Include students in schools that do not have a transitional-first grade class to determine if placement makes a difference.

3. Use a more refined level of measurement on Gesell tests.

4. Administer the IQ measure nearer in time to the Gesell test.
APPENDIX

DISTRIBUTION OF THE ACHIEVEMENT TEST SCORES FOR GESELL SUBTESTS
Figure 1. Distribution of the Word Attack Scores for the Gesell Cubes Subtest.
Figure 2. Distribution of the Vocabulary Scores for the Gesell Cubes Subtest.
Figure 3. Distribution of the Reading Comprehension Scores for the Gesell Cubes Subtest.
Figure 4. Distribution of the Language Scores for the Gesell Cubes Subtest.
Figure 5. Distribution of the Word Attack Scores for the Gesell Name Subtest.

65 or Below
N = 22
Mean: 374
Standard Deviation: 113

66 or Above
N = 50
Mean: 475
Standard Deviation: 115
Figure 6. Distribution of the Vocabulary Scores for Gesell Name Subtest.
Figure 7. Distribution of the Reading Comprehension Scores for Gesell Name Subtest.

- 65 or Below
  - Mean: 397
  - Standard Deviation: 136

- 66 or Above
  - Mean: 419
  - Standard Deviation: 132

<table>
<thead>
<tr>
<th>Score Range</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>200-224</td>
<td>1</td>
</tr>
<tr>
<td>225-249</td>
<td>2</td>
</tr>
<tr>
<td>250-274</td>
<td>2</td>
</tr>
<tr>
<td>275-299</td>
<td>3</td>
</tr>
<tr>
<td>300-324</td>
<td>3</td>
</tr>
<tr>
<td>325-349</td>
<td>3</td>
</tr>
<tr>
<td>350-374</td>
<td>3</td>
</tr>
<tr>
<td>375-399</td>
<td>3</td>
</tr>
<tr>
<td>400-424</td>
<td>4</td>
</tr>
<tr>
<td>425-449</td>
<td>6</td>
</tr>
<tr>
<td>450-474</td>
<td>6</td>
</tr>
<tr>
<td>475-499</td>
<td>5</td>
</tr>
<tr>
<td>500-524</td>
<td>5</td>
</tr>
<tr>
<td>525-549</td>
<td>3</td>
</tr>
<tr>
<td>550-574</td>
<td>2</td>
</tr>
<tr>
<td>575-599</td>
<td>1</td>
</tr>
</tbody>
</table>

N = 10

Mean: 39

Standard Deviation: 136

Mean: 42

Standard Deviation: 132
Figure 8. Distribution of the Language Scores for Gesell Name Subtest.
Figure 9. Distribution of the Word Attack Scores for Gesell Copy Forms Subtest.

65 or Below
N = 27
Mean: 411
Standard Deviation: 118

66 or Above
N = 45
Mean: 460
Standard Deviation: 123
Figure 10. Distribution of the Vocabulary Scores for Gesell Copy Forms Subtest.
Figure 11. Distribution of the Reading Comprehension Scores for the Gesell Copy Forms Subtest.

65 or Below
N = 17
Mean: 379
Standard Deviation: 135

66 or Above
N = 35
Mean: 426
Standard Deviation: 128
Figure 12. Distribution of the Language Scores for the Gesell Copy Form Subtest.

- 65 or Below: N = 27, Mean: 364
- 66 or Above: N = 45, Mean: 409

Standard Deviation: 121 for 65 or Below
Standard Deviation: 128 for 66 or Above
65 or Below
N = 21
Mean: 382
Standard Deviation: 123

66 or Above
N = 51
Mean: 467
Standard Deviation: 114

Figure 13. Distribution of the Word Attack Scores for the Gesell Numbers subtest.
65 or Below  
N = 21  
Mean: 374  
Standard Deviation: 87

66 or Above  
N = 51  
Mean: 456  
Standard Deviation: 87

Figure 14. Distribution of the Vocabulary Scores for the Gesell Numbers Subtest.
Figure 15. Distribution of the Reading Comprehension Scores for the Gesell Numbers Subtest.
Figure 16. Distribution of the Language Scores for the Gesell Numbers Subtest.
65 or Below
N = 33
Mean: 398
Standard Deviation: 124

66 or Above
N = 39
Mean: 480
Standard Deviation: 109

Figure 17. Distribution of the Word Attack Scores for the Gesell Incomplete Man Subtest.
Figure 18. Distribution of the Vocabulary Scores for the Gesell Incomplete Man Subtest.

65 or Below
N = 33
Mean: 413
Standard Deviation: 93

66 or Above
N = 39
Mean: 447
Standard Deviation: 94
Figure 19. Distribution of the Reading Comprehension Scores for the Gesell Incomplete Man Subtest.

65 or Below
N = 21
Mean: 381
Standard Deviation: 151

66 or Above
N = 31
Mean: 433
Standard Deviation: 112
65 or Below
N = 33
Mean: 366
Standard Deviation: 131

66 or Above
N = 39
Mean: 416
Standard Deviation: 120

Figure 20. Distribution of the Language Scores for the Gesell Incomplete Man Subtest.
Figure 21. Distribution of the Word Attack Scores for the Gesell Interview Subtest.

- 65 or Below
  - N = 25
  - Mean: 360
  - Standard Deviation: 104

- 66 or Above
  - N = 47
  - Mean: 489
  - Standard Deviation: 108

600/over
575-599
550-574
525-549
475-499
450-474
425-449
400-424
375-399
350-374
325-349
300-324
275-299
250-274
225-249
200-224
199/under
Figure 22. Distribution of the Vocabulary Scores for the Gesell Interview Subtest.
65 or Below
N = 13
Mean: 321
Standard Deviation: 153

66 or Above
N = 39
Mean: 446
Standard Deviation: 109

Figure 23. Distribution of the Reading Comprehension Scores for the Gesell Interview Subtest.
Figure 24. Distribution of the Language Scores for the Gesell Interview Subtest.

65 or Below
N = 25
Mean: 312
Standard Deviation: 99

66 or Above
N = 47
Mean: 440
Standard Deviation: 121
65 or Below
N = 26
Mean: 391
Standard Deviation: 110

66 or Above
N = 46
Mean: 474
Standard Deviation: 120

Figure 25. Distribution of the Word Attack Scores for the Gesell Animals Subtest.
Figure 26. Distribution of the Vocabulary Scores for the Gesell Animals Subtest.
Figure 27. Distribution of the Reading Comprehension Scores for the Gesell Animals Subtest.
Figure 28. Distribution of the Language Scores for the Gesell Animals Subtest.
65 or Below  
N = 28  
Mean: 414  
Standard Deviation: 129

66 or Above  
N = 44  
Mean: 462  
Standard Deviation: 116

Figure 29. Distribution of the Word Attack Scores for the Gesell Interest Subtest.
Figure 30. Distribution of the Vocabulary Scores for the Gesell Interest Subtest.
Figure 31. Distribution of the Reading Comprehension Scores for the Gesell Interest Subtest.
Figure 32. Distribution of the Language Scores for the Gesell Interest Subtest.

65 or Below
N = 28
Mean: 362
Standard Deviation: 133

66 or Above
N = 44
Mean: 417
Standard Deviation: 122
REFERENCES AND NOTES


