THE DEVELOPMENT AND EVALUATION OF THE PREKINDERGARTEN READINESS SCREENING DEVICE

By

Richard Norman Claus

A DISSERTATION

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

DOCTOR OF PHILOSOPHY

Department of Educational Psychology

1981

ABSTRACT

THE DEVELOPMENT AND EVALUATION OF THE PREKINDERGARTEN READINESS SCREENING DEVICE

By

Richard Norman Claus

The Prekindergarten Readiness Screening Device (PRSD) is an individually administered 27 item applied performance checklist. Scores from PRSD have been used to determine the eligibility for participation in a Title I Prekindergarten program. The purposes of this research were: (a) to estimate the rater reliability, (b) to explore age progression of items as one aspect of construct validity, and (c) to estimate predictive validity of the PRSD.

Subjects for the rater reliability study were the 11 teachers and 8 aides of Saginaw Public School's Prekindergarten program. This study consisted of raters scoring two videotaped testing sessions. Intra-class correlation obtained for an individual was .789 and for the average of all raters was .973.

Subjects for the construct validity study were the 1,415 potential prekindergarteners that resided in the economically depressed area of Saginaw, Michigan designated as the Title I attendance areas. These subjects were screened for either the school year of 1978-79, 1979-80, or

1980-81. The screening data were used to study the age progression of each item by means of a test for trend in order chi-square contingency tables. The null hypothesis of no age progression by age interval in the percent passing each item was not accepted for 25 of 27 items (alpha < .05). On the basis of Cramer's contingency coefficients related to the trend of the 25 items showing significant results (an average coefficient of .136) and a Pearson correlation coefficient between age and total PRSD score (r = .161), percentile norms were calculated for the entire group rather than for each age group. The data set also yielded reliabilities of .820 and .860 for the Kuder-Richardson 20 and Spearman Brown split-half formulas respectively.

Subjects for the predictive validity study were the 396 potential prekindergarteners screened with the PRSD during the 1978-79 school year. These subjects were subsequently, during the same school year, pre- and post-tested with the Prekindergarten Saginaw Objective Referenced Test (PSORT) and tested with the Metropolitan Achievement Test (MAT) in April of the 1979-80 school year. The Pearson correlation coefficients between the PRSD and each achievement instrument follow: pre-test PSORT, r = .487; post-test PSORT, r = .383; and MAT, r = .484. Fisher's r to Z transformation of the correlations showed all three relationships to be significantly different than zero (alpha < .05).

ACKNOWLEDGMENTS

Many people have helped me complete this research. I would like to acknowledge their names with deep appreciation and beg the indulgence of those I may overlook but equally thank.

First to Dr. Robert L. Ebel, my advisor and committee chairman, for his personal interest, guidance and encouragement throughout the doctorate program and this research. I would also like to thank the other members of my dissertation committee Dr. Andrew C. Porter, Dr. Samuel A. Moore II, and Dr. Philip M. Marcus for their valuable insights and contributions to this study and their guidance and inspiration during the course of the author's graduate career.

The author is also indebted to Mrs. Reva Ruby and her staff for their participation in the data collection procedures of the Title I Prekindergarten program. Their interest and cooperation in always trying to provide a better prekindergarten program made this study possible.

My sincere thanks to my fellow colleagues at the Saginaw Public Schools, Mr. Barry Quimper, Mr. Michael Manley, Mr. David Lutenski, and Mr. Hsing Chen for their friendship, inspiration, support, and editorial comments.

I am especially grateful to Mr. Hsing Chen for aiding me in the data processing aspects of the study.

My manuscript was typed by Miss Terry Fisher, who labored many hours to type the many drafts. I owe her a special debt of gratitude.

I am appreciative to my mother and father, Mrs. Dorothy Claus and Dr. Norman Claus for their belief in the value of education, continuous encouragement, and financial support. To them goes a very personal and special thanks that words cannot express adequately.

I am thankful to my wife Sachiko, who has been a continuous source of motivation and inspiration for me to complete my doctoral program. Her constant faith in me, her personal sacrifice, and her loving, unfaltering support made everything possible.

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont'd.)

TABLE OF CONTENTS (Cont'd.)

TABLE OF CONTENTS (Cont'd.)

TABLE OF CONTENTS (Cont'd.)

LIST OF TABLES

LIST OF TABLES (Cont'd.)

LIST OF TABLES (Cont'd.)

LIST OF FIGURES

CHAPTER 1: THE PROBLEM

This study deals with the early identification of children who may have problems in school. A survey of psychological and educational literature of the recent past (Lesiak, 1978) indicates a growing concern with this problem. A great deal of activity has taken place, specifically in the past decade and a half, in the construction of standardized school readiness instruments. This interest in early detection devices is clearly evident in the upward trend in the number of screening measures reviewed in the Mental Measurement Yearbook (Buros, 1965, 1972, 1978). The sixth edition (1965) of the Mental Measurement Yearbook (MMY) mentioned eight readiness tests, compared with 29 in the seventh edition (1972) and 16 in the eighth edition (1978). While the number of new instruments appearing in the latest MMY shows a decline (16 as opposed to 29), considerable interest still seems evident and, as the next section relates, adequate school readiness screening devices are needed.

NEED

This section reviews issues pertaining to the need for readiness and screening instruments. The topics specifically covered are the following: inadequacies of available tests, effectiveness of preschool programs, and the need for better prekindergarten readiness instruments.

Inadequacies of Available Tests

The most disheartening observation to a psychometrically inclined reviewer of the Mental Measurement Yearbook critiques of readiness and screening measures is the scarcity of data set forth as essential for test development by the American Psychological Association. Of the 29 readiness tests reviewed in 1972, 11 (37.9%) lacked at least one of the following: norms, description of standardization sample, data on test reliability, or data on test validity. An additional five tests gave only partial information on an essential dimension (Maitland, Nadeau, and Nadeau, 1974; Buros, 1972). Of the 16 tests reviewed in 1978, 9 (56.2%) lacked data on at least one of the essential dimensions listed above.

When screening and evaluating the readiness of a special population is a major concern, then the number of satisfactory instruments are few or nonexistent. This is just the problem with the screening and evaluating of disadvantaged children for enrollment in Title I

Prekindergarten and Head Start Programs throughout this country. Assessment processes for these federally funded programs have been complicated by the absence of suitable and adequate measurement instruments (Grotberg, 1969; Walker, 1972). Even though a select number of adequate instruments may exist, the time available for individually screening each child is usually limited. Of the 16 tests reviewed in the Eighth Edition of the Mental Measurement Yearbook, the median time required for administration was 32.5 minutes (Buros, 1978) with only three instruments taking less than 15 minutes (see Appendix A for times on all instruments). The identification process appears to be a problem with some prekindergarten staffs because they see teaching as their primary responsibility and the screening process as secondary. A shorter reliable and valid screening instrument would help alleviate this problem.

In the Saginaw Public Schools Title I program, a need has existed for many years for both screening and outcome instruments normed on a population that included representatives of non-white segments of the population. Having a large non-white segment in the Saginaw Title I population, it was both scientifically sound and politically astute to request such norming samples of test publishers. However, the search for both sound preschool instruments and tests normed on the appropriate population was unproductive.

Approximately four years ago a promising outcome measure was found (namely the Iowa Test of Preschool Development), but unfortunately it was normed on a rural all white population. The author of the test explained that the all white sample provided a goal for non-whites. (See Appendix B for copies of correspondence on this matter.) Norms of a meaningful referent group are essential to describing the meaning of scores to parents. If the skills measured are developmental in nature as most pre-kindergarten screening tests claim, then norms can also provide further proof of age differentiation and construct validity (Anastasi, 1970, p. 474). Whitely and Davis (1974, pp. 163-178) explain that, under classical models of test development, the score obtained by a person is not interpretable without referring to both some norm group and the particular test forms used. Norms seem essential to the further development of any measurement instrument and to the practical task of explaining scores to parents.

Effectiveness of Preschool Programs

Data available from the preschool efforts of the 1960's and early 1970's indicate that lasting positive effects and academic gains are possible through preschool programs aimed at disadvantaged youngsters. The study of lasting effects of preschool by Lazar and

Darlington (1979) found the following effects of preschool programs: fewer children assigned to special education classes regardless of their initial abilities or early home background; fewer children retained in grade (again regardless of initial abilities or home background); significantly increased children's scores on fourth grade (grade level at which the most data were available) mathematics achievement tests with a suggestive trend toward increased scores on fourth grade reading tests; higher I.Q. scores than control children up to age 12; and increased likelihood to give achievement-related reasons for being proud of themselves.

Need for Better Readiness Instruments

A number of factors, taken together, indicate that more valid and reliable preschool readiness instruments will be demanded in the 1980's. Some of these factors are declining student enrollment, excess numbers of teachers in the work force, and proven preschool programs. Declining student enrollments seem likely to continue, resulting in an over supply of teachers seeking employment. Faced with this surplus teacher work force educators and planners have attempted to expand the scope of education at both ends of the learning continuum. Much talk centers upon the potential preschool aged pupils and adult continuing education stduents as populations where new programs need to be

developed or existing programs need to be expanded. However, should the economic conditions of the 1980's continue, this factor will make program expansion more likely in an area where there has been a proven need. A number of prekindergarten programs do appear to satisfy a need. However, it would seem that program expansion would be limited to those programs that have established a need based upon valid and reliable screening instruments.

On the other hand, if resources for education do continue to be reduced, then the size of existing preschool programs will be cut. Thus better screening of possible preschool aged participants to find those most lacking school readiness skills would still be a necessity. Yet normed, valid, and reliable screening devices for disadvantaged minority preschool aged population are not presently available.

PURPOSE

Th

develop

instrum

instrum

of deve

able to

Further

prekind

such pr

teristi

PURPOSE

The specific purpose of this study is to further develop and evaluate a teacher administered measurement instrument for screening prekindergarten pupils. The instrument would be used to identify children, who because of developmental and/or experiential problems may be less able to meet the typical expectations of their peers. Further, this instrument should satisfy concerns of Title I prekindergarten staff. A screening instrument used for such programs ideally should have the following characteristics:

- An administration length of no more than 12 minutes.
- Easily understood directions so that teachers and paraprofessionals could administer the test after a 90 minute training session.
- Standardized administration and scoring procedures to ensure uniform results.
- A review of pupil performance in an applied setting.[1]

---

[1]In other words, program personnel want to observe requested pupil behaviors in a school setting (applied setting). Other possible screening measures as the Vineland Social Maturity Scale call for a parental report of recalled child performance in a home setting. Thus the desire of our Title I Prekindergarten supervisor and staff is to obtain a screening measure that focuses on actual child performance in a non-home situation. Both the Title I personnel and this researcher believe this type of obtained data is more reliable and valid than what parents can recall from memory. In addition, the strange environment of the schools adds a dimension of realism not unlike what a child would face in school. Terming this test characteristic as a requirement reemphasizes the applied performance setting that did serve to focus our developmental efforts.

Af

this re

all the

Th

prekind

qualiti

n more

greates

C

instru

- Normative data from an E.S.E.A. Title I population with a large minority racial-ethnic segment.
- A sampling of school readiness skills[1] appropriate for three, four, and five year olds.
- Validity and reliability data on the appropriate norming group.

After reviewing several books[2] of test critiques, this researcher is still unable to find a test that meets all the requirements stated above.

The dissertation would result in developing a shorter prekindergarten screening test with known psychometric qualities and norms. The ultimate benefit is the ability to more accurately assess and select those pupils in greatest need of prekindergarten services.

Other benefits which might be derived from such an instrument are:

- Lower administrative costs resulting from tests administered by teachers and/or aides not to mention the shortened amount of time to administer. (This would negate the necessity to hire school psychologists.)

---

[1]School readiness skills relate to achievement in gross motor, fine motor, language, cognitive, and personal social development.

[2]The Mental Measurement Yearbooks by Oscar Buros; Tests and Measurements in Child Development: A Handbook by Orval Johnson and James Bommarito; Handbook for Measurement and Evaluation in Early Childhood Education by William Goodwin and Laura Driscoll; and CSE-ECRC-Preschool Kindergarten Test Evaluations edited by Ralph Hoepfner, Carolyn Stern and Susan Nummedal.

dete

cons

norm

the

inst

- Local norms applicable to other Title I prekindergarten populations with similar racial-ethnic compositions.
- Reliability estimates to assess the consistency and replicability of results.
- Chronological age norms to demonstrate the developmental nature of the school readiness skills chosen (construct validity).
- Predictive validity estimates to allow the prospective user to judge this instrument's validity against other instruments.

In summary, the purposes of this study are: (a) to determine the reliability, (b) to obtain estimates of construct and predictive validity, and (c) to develop norms to make the results interpretable in comparison to the norming group of a prekindergarten readiness screening instrument.

DEFINITIO

The

common m

Applied

App

which th

surround

situatio

1980, p.

there ar

ing that

Prekinde

Pre

as young

cannot b

Prekinde

Pre

administ

Standard

acceptab

assessed

determin

Prekinde

DEFINITION OF TERMS

The following definitions are supplied to provide a common meaning to key concepts in this study.

Applied Performance Tests

Applied performance tests refer to instruments in which the test stimulus, the desired response, and the surrounding conditions approximate the reality of an actual situation drawn from a specific role-based context (Slatter, 1980, p. 2). As can be inferred in the word "approximate" there are many alternative approaches to performance testing that range on the scale of realism.

Prekindergarten Aged Children

Prekindergarten aged children are defined in Michigan as youngsters who must be four years old by December 1 and cannot be five years old by December 2.

Prekindergarten Readiness Screening Device (PRSD)

PRSD (see Appendix C for copy) is an individually administered 27 item applied performance checklist with standardized spoken directions and a statement of an acceptable response for each item. The instrument assesses entry behavior and provides information for determining the eligibility for participation in Title I Prekindergarten programs.

School R

Sch

a given

ing atte

readines

School Readiness

School readiness refers to the ability to engage in a given school activity depending on the learner's existing attention set, motivation, and state of developmental readiness (Gagne, 1970, pp. 277-301).

TEST EVAL

The

the state

PRSD.

TEST EVALUATION QUESTIONS

The following questions were formulated in light of the stated purpose to focus the further evaluation of the PRSD.

(1) What is the estimated rater reliability of the PRSD?

(2) How does the ability of male and female prekindergarten aged pupils vary on the PRSD as a function of chronological age?

(3) What is the strength of the relationship between school readiness and future measures of school achievement?

OVERVIE

The

tion is

first se

in the

the pre

ments p

sound c

the des

qualiti

(PRSD).

cussed.

offered

OVERVIEW

The general plan for the remainder of this dissertation is as follows. Chapter 2 deals with two topics. The first section contains a review of the historical changes in the mission of preschool education from the 1920's to the present. The second section deals with the key elements pertaining to the construction and use of technically sound checklists. Chapter 3 consists of a description of the design and methods used to evaluate the psychometric qualities of the Prekindergarten Readiness Screening Device (PRSD). In Chapter 4 the findings are presented and discussed. In Chapter 5, the summary and conclusions are offered plus recommendations for future studies.

The

sections

in the mi

objective

the expan

preschool

of const

to presch

this sec

list ass

CHAPTER 2: REVIEW OF LITERATURE

The review of literature is divided into two major sections. The first section reviews the historical changes in the mission of early childhood education. The primary objective of this portion is to provide basic knowledge of the expanded range of educational objectives of present day preschool programs. The second section reviews principles of construction and administration of checklists as related to preschool identification. The primary objective of this section is to review some of the difficulties in checklist assessment.

REVIEW O

Thi

childre

childhoo

introduc

educatio

The

the peri

ized pro

general

is offe

The revi

of early

review o

and thos

emergen

childho

differe

Hopefu

necessa

school

Histor

Th

were cr

Serious

REVIEW OF LITERATURE - EARLY CHILDHOOD EDUCATION

This review of literature relates to education of children under the age of six, henceforth termed early childhood education. It is undertaken to provide an introduction to historical practices in early childhood education.

The time frame of the review is primarily devoted to the period from the 1920's to the present when most organized programs in the United States took place. A brief general review of early childhood education prior to 1920 is offered to provide the context for later developments. The review is focused on the following subtopics: history of early childhood education prior to 1920; historical review of differences between objectives of nursery schools and those of preschools; historical factors leading to the emergence of innovative programs of the decade of early childhood education (1965-1975); and generalization and differences of the early childhood programs of this decade. Hopefully, this review will impart an appreciation of the necessary breadth of skills embodied in present-day preschool programs.

History of Early Childhood Education Prior to 1920

Until the twentieth century, the years of childhood were on the whole, a most unhappy period of existence. Serious investigators of ancient child rearing practices

report

the am

to be

der of

centur

did re

ninete

as a m

during

quotes

"Withh

him wit

the roc

the eig

spoil t

battere

In

in the

Harping

summari

report that the further one goes back in history the less the amount of child care and the more likely children were to be abandoned, beaten, terrorized and abused. The murder of babies or infanticide was a common practice for centuries (Osborn, 1975). While the beginning of Christianity did reduce this practice, infanticide persisted into the nineteenth century. The western world's view of the child as a miniature adult, allowed children to be exploited during the Industrial Revolution. Osborn (1975, p. 12) quotes The Harmony of the Gospels printed in 1678 as follows: "Withhold not correction from the child, for if thou beatest him with the rod he will not die, Thou shall beat him with the rod and deliver his soul from hell." Thus prior to the eighteenth century, the attitude of spare the rod and spoil the child caused many children to be, in reality, battered and abused.

In this country the 1860-1930 period marked a rise in the childhood population which resulted in an accompanying change in attitudes toward childhood. Osborn (1975) summarized this period as follows:

Histor

C

over t

instru

school

1960's

day ca

socio-

1971;

and Se

howeve

tion b

taged"

215; I

Goodla

71; St

school

> There was a rise in child population from 17 million in 1860 to 48 million by 1930. However, in percentage of total population there was a sharp decline (51% to 38%). Thus, as children became more visible, their needs became more apparent. In addition, more adults were available to devote time to care and rear children. These factors, coupled with the rising opposition toward child labor and the increasing awareness of the need for education, helped to change attitudes concerning the role of the child in society (p. 22).

Historical Review of Objectives of Early Childhood Programs

Contrasted with the history of education for children over the age of six, the account of nursery and preschool instruction has been brief. In the United States nursery school instruction began in the 1920's. Until the early 1960's the goal of nursery schools was mainly to furnish day care for working-class children and to encourage the socio-emotional growth of middle-class children (Cazden, 1971; Dowley, 1969; Kamii, 1971; Osborn, 1975, pp. 38-61; and Sears and Dowley, 1963). By the middle of the 1960's, however, the bulk of the intended recipient of such education had shifted to the lower-lower-class, or "disadvantaged" and was now termed preschool (Fallon, 1973, pp. 207-215; Lazerson, 1971; Goodwin and Driscoll, 1980, pp. 3-6; Goodlad, Klein, and Novotney, 1973; Osborn, 1975, pp. 61-71; Stanley, 1972; and Shane, 1971). As the terms "nursery school" and "preschool" suggested, the focus of the

educational effort had changed. The term nursery school has always tended to imply a downward extension of the functions related to the family. While preschool has little relationship to the family in its derivation, preschool suggested a downward extension of the school to prevent the failure of "disadvantaged" children in school.

A short historical review may clarify the difference between the curricular objectives of the nursery schools and those of the later preschools.

The majority of the first nursery schools in the United States were established by colleges and universities for research purposes. The research purposes were related to the discovery and demonstration of "better" ways of caring for young children. The objectives of these nursery schools varied according to whether the sponsoring department of higher education was home economics, psychology, medicine, or education. The curricular objectives of most of these schools consisted, however, primarily of habit training, for example, eating, napping, and the promotion of physical health (Sears and Dowley, 1963, p. 815).

The first involvement of the federal government in nursery school education occurred during the depression years of the 1930's under the Works Progress Administration (WPA) program. The primary objective of the WPA program was to provide work for unemployed teachers, nurses and

helpers from 1933 to 1942 (Osborn, 1975, p. 48). The curriculum continued to stress physical health and the "good" habits of sleeping, elimination, dressing, washing, and eating, and so forth, with the surplus food made available from the government program of economic supports for farmers (Kamii, 1971, p. 283).

The World War II years brought about another spurt of growth for the nursery school movement. The federal funds of the Lanham Act provided for day care centers for young children whose mothers worked in strategic war industries. Although the curriculum continued to emphasize habits and routines related to health and welfare of the child with longer period of "free" play, a new concern for childrens emotional well being became more widespread as the war related child care programs continued (Sears and Dowley, 1963, p. 815). This new concern for youngsters' social-emotional life came from a number of sources including: Freudian theory; longer hours at nursery school resulting in more behavior problems; disturbances in parent-child relationships as the result of the mother's employment and the father's absence (Stolz, 1954); the writings of Frank (1938), Gesell (1940, 1943), and Spock (1946). Frank analyzed the essential learnings of early childhood in relation to the child's fundamental emotional needs as a feeling, responding individual. He cautioned teachers about

dangers

of chil

display

develop

repeata

years o

cance o

Spock a

and lea

the chi

as child

function

p. 285).

The

for nurs

who had

their f

the case

school n

cooperat

the nurs

and emot

providin

Sent pla

traditio

in the

dangers of excessive expectations resulting in disturbances of children's personal, social, and cultural growth. Gesell displayed hundreds of physiological and mental aspects of development which he believed, followed a more or less repeatable pattern in all children during the first five years of life. He, in particular, emphasized the significance of the developmental elements in emotional development. Spock assigned new significance to the concept that growing and learning proceed more smoothly if allowed to occur in the child's own way and time. While nursery schools began as child care centers, they thus gradually took on the function of providing preventive psychiatry (Kamii, 1971, p. 285).

The post-war years saw the stoppage of federal funds for nursery school care based on the assumption that women who had been drawn into wartime industries would return to their full-time home responsibilities. This was not always the case, however, and the parent cooperative nursery school movement sought a variety of ways and means through cooperative efforts to keep child care centers open. While the nursery school movement had fostered children's social and emotional growth ever since its beginning as part of providing good care, the cooperative nursery school movement placed a new conscious effort on encouraging socio-emotional growth. This objective came to be expanded upon in the 1940's (Dowley, 1969, pp. 320-321).

A

develo

since

emphas

contri

progra

ness a

the ci

ways t

land.

fundin

more s

ment o

childre

whole,

focusir

were de

opmenta

public

screens

child's

if plac

was the

descript

satory p

Although as Kamii (1971, p. 285) found that cognitive development was a concern of many nursery school educators since the 1920's, it was not until the 1960's that real emphasis was placed on it. In the 1960's two major factors contributed to a shift from nursery school to preschool programs that emphasized cognitive growth and other readiness abilities necessary for success in school. One was the civil rights movement that, among other things, sought ways to provide more opportunities to minorities of this land. The answer from the federal government was massive fundings for both lower and middle class youngsters. A more striking result, already alluded to, was the establishment of a large number of new educational programs for children in the preschool or early school years. On the whole, these new programs had a "compensatory" flavor; focusing primarily on children from poor families, they were designed to remedy the usually one to three year developmental lag of the entering disadvantaged children in the public school setting. A vast majority of these programs screened children so that their staffs could determine the child's educational needs before school entry and determine if placement in their program or some other special program was the best decision (Zeitlin, 1976). Fallon (1973) described the general theory that most of these compensatory preschool programs operate under as:

> How well and how rapidly children develop their mental model of the world depends largely on their environment. The more the child has seen and heard, the greater is his desire to see and to hear. The greater the variety of things he has learned to cope with, the greater his capacity to cope. Much that traditionally has been taught to older children can and should be taught in the early years (p. 208).

This brief historical review has shown the evolution of the nursery school as starting with general objectives that became more and more differentiated. This evolution came about as a result of social and historical forces that focused on different aspects of the developing child.[1] In reality, it has never been possible to separate the physical care of the children (emphasis of traditional nursery schools since 1920's) from their socio-emotional development (emphasis of cooperative parent nursery schools since 1940's) or their socio-emotional development from intellectual and other school readiness abilities development (emphasis in preschools since 1960's).

---

[1]These forces will be explored in greater detail in the next section.

The various preschool programs in existence today differ considerably in their objectives and methods of instruction as evident by a study of various prekindergarten programs (Cazden, 1971; Chapman and Lazar, 1971; Fallon, 1973; Goodwin and Driscoll, 1980, pp. 421-468; Hess and Bear, 1968; Kamii, 1971; National School Public Relations Association, 1970, 1973; Sears and Dowley, 1963; and Stanley, 1972). These references capture the range of nursery school and preschool objectives from the 1920's through the late 1970's. Kamii (1971, p. 286) offered the following broad objectives of preschool education:

1. Socio-emotional development,
2. Perceptual-motor development,
3. Cognitive development, and
4. Language development.

A review of the above sources bear out that these categories still were appropriate through the late 1970's as a means to categorize broadly early childhood objectives.

Factors Related to the Renewed Interest in the 1960's of Early Childhood Education

In the United States in the 1960's, a meeting of academic and socio-political forces created productive conditions for a rebirth of activity in early childhood education (Dowley, 1969, 1971; Fallon, 1973; Goodwin and Driscoll, 1980, p. 3; Osborn, 1975, pp. 54-55; Shane, 1971). A succession of persuasive statements from the academic world

asse

Kirk

reta

cati

grow

Brun

Intel

taug

any c

Hart

fixed

appo

Alle

oper

cont

Sugg

Supp

Incre

the s

Studi

A sec

the c

Incen

Stimu

At p

asserted very strongly the importance of early experience. Kirk's experimental work (1958) with educable mentally retarded children showed that the enriched preschool educational environment, on the whole, increased rates of growth following educational opportunities at a young age. Bruner emphasized the role of education in children's intellectual development by stating, "Any subject can be taught effectively in some intellectually honest form to any child at any stage of development" (1960, p. 33). Hunt (1961) questioned the notion that intelligence is fixed at birth and contended that an enriched environment, especially in early childhood, could make a meaningful difference in both the rate and level of intellectual development. Bennett, Diamond, Kretch, and Rosenzweig (1964) conducted research on infrahuman subjects that strongly suggested that memory cells, brain size, and the blood supplied to the cerebral hemispheres actually can be increased by changes in the environment to create stimulating surroundings. Bloom's research (1964) of longitudinal studies determined that intellectual development occurs at an accelerated rate in the early years of life and therefore concluded that environment is most critical then. The concern for mental development was further aroused by a growing recognition of Piaget's work. Hall (1965, 1966) and other cultural anthropologists accumulated suggestive

evidence that it is during the first four or five years of life that many personal behaviors (such as language, attitude, values, and ways of learning) begin to take on the form they will retain for a lifetime. All these works emphasized the importance of the early childhood years as the foundation for later ability and development.

Various socio-political forces also produced interest in early childhood education. These forces undoubtedly served as a greater stimulus to action than the academic forces. The civil rights movement and 1954 Supreme Court decision (i.e., striking down the concept of separate but equal schools for each race) helped focus on the needs of minorities. The violence and social unrest of the early 1960's in many urban areas were thought to be rooted in the poverty-disadvantaged cycle - a frustrating, devastating treadmill, which particularly trapped young, poor black and other minorities (Goodwin and Driscoll, 1980, p. 3). A large number of minority youngsters entered school behind in achievement and fell behind further each year in school. Such events and theory led, at least in part, to the federal war on poverty as well as attempts to equalize educational opportunity through such efforts as Headstart, Home Start, Title I of the Elementary and Secondary Act of 1965, and school integration. During the same time the rising divorce rate plus the increasing

number of women entering the work force emphasized the need for child care. The women's liberation movement advanced child care as a high priority need for women in general. The "Sputnik Counterreaction" served as another socio-political force for the advancement of early childhood education. After the jolting realization of the Soviet advancement in space exploration provided by Sputnik in 1957 and the early wave of criticism of American education, there seemed to be in the early 1960's a reaffirmation of public support for education.

One result of these converging factors was a gradual increase in preschool enrollment, especially among children of the middle class, and an increase in day care services of programs to help prevent school failures among the lower-lower-class children. Another was the accumulation of knowledge suggesting that the early childhood years were the most susceptible to intervention.

Federal fundings for Headstart and Title I Preschool programs still continued into the 1980's. Lazar and Darlington (1979), as reported in Chapter One in greater detail, found in their longitudinal study that lasting positive effects and academic gains were possible through preschools aimed at disadvantaged youngsters in cognitive oriented programs. Other recent newspaper articles ("Preschool Education Pays Off for Students, Study Says,"

1980; "Preschool Research Shows Lasting Benefit," 1980; "Preschool Programs Cost Effective, Riles Says," 1980) reported that preschool programs have been somewhat effective in showing some lasting benefits, but failed to wipe out the differences between disadvantaged and non-disadvantaged youngsters as once promised. Still other newspaper articles as "The Ups and Downs of Preschool Intervention" (1980) reported that early childhood is still being oversold. This article suggests that researchers would be better off entertaining the hypothesis that there are certain environmental nutrients that optimize development of all stages.

Generalizations and Differences of Present-Day Early Childhood Programs

Osborn (1975, pp. 62-63) offered a number of generalizations and differences stemming from an analysis of the wide variety of innovative early childhood programs developed between 1965-1975. Most of these programs were still in operation as the 1980's began with the prospects of increasing numbers of preschoolers taking part in these activities (Goodwin and Driscoll, 1980, pp. 3-6). Most developers of preschool programs adhered to the following generalizations in the operation of their programs.

1. Children should not be left to "unfold" in accordance with the nature of the child and the natural environment.

2. Children under the age of six were expected to gain from an enriched and systematic curriculum.

3. Involvement of parents in the homework assignments and activities was essential.

4. The enhancement of the self-image (or ego development) of the learner was essential.

5. Teachers allowed time to provide for individual instruction.

6. Skills emphasized were deemed to insure success in later academic performance.

Although all programs have emphasized the importance of the early childhood years, some programs have stressed education before two years of age mainly through home instruction. Program models also differed along the following dimensions.

1. Structure: Some programs used a great deal of structure (e.g., Engelmann-Bereiter Program); some a moderate amount (e.g., Montessori Schools); and others very little (e.g., British Infant school). Fallon (1973) provided a good synopsis of each of these specific programs mentioned above plus examples for the remaining dimensions listed below.

2. Reinforcement: While all models strove to ultimately rely on intrinsic reinforcement, some programs used token reinforcements and other programs stressed social reinforcement.

3. Curriculum: Some programs focused on cognitive aspects of the curriculum; others emphasized social-emotional, creative arts. It must be noted that all programs recognized the importance of language.

4. Teacher Role: Some programs stressed a passive non-directive role for the teacher; others emphasized the teacher's role as the "stage setter;" and others accented the active, direct role of the teacher.

5. Activities: In some programs, children have a large assortment of activities available, while in others there is little or no choice in activity selection.

Osborn (1975) ends these observations with the following:

> Perhaps the final generalization is the most valuable to remember--To date no single program has been found to be the best program for all children. All models can "point with pride" at their successes and "view with alarm" some of their shortcomings (p. 64).

REVIEW
TRAFFIC

I

school

are co

school

sters

ing sul

rating

sources

means t

Assumpt

Cu

Ch

quanti

scales

choice)

tion or

require

the rat

Since

REVIEW OF LITERATURE - CHECKLIST CONSTRUCTION AND ADMINISTRATION

In an attempt to discover which youngsters need preschool programs, some knowledge about how rating scales are constructed and can be employed to reliably assess school readiness of potential compensatory education youngsters seems essential. The review consists of the following subtopics: assumptions behind the construction of rating scales generally and checklists in particular, sources of common errors associated with checklists, and means to overcome these errors.

Assumption Behind the Construction of Checklists

Guilford (1954) stated,

> The use of ratings rests on the assumption that the human observer is a good instrument of quantitative observation, that he is capable of some degree of precision and some degree of objectivity. His ratings are taken to mean something accurate about certain aspects of the object rated (p. 278).

Checklists have been described as unique in terms of quantitative judgment compared with other forms of rating scales (i.e., numerical, graphic, standard, and forced choice) because checklists require the least discrimination on the part of the rater. Specifically, the rater is required to use a two step scale. Thus only cases near the rater's threshold present difficulty in judging. Since scoring of cumulative point checklists usually

involve weights of +1 or 0 for each item, scoring was also very simple.

Guilford (1954, p. 273) speaking of checklists specifically listed the applications in the following areas: employee's value to his/her organization, simple traits of personality, achievement tests, and proficiency measures. Guilford (1954) stated,

> When the items are of specific actions that are observed by the rater, the checklist becomes essentially an achievement or proficiency test and its score has the status that would be accorded to that type of measure (pp. 273-274).

Source of Common Errors Associated with Checklists

Checklists have a number of possible common sources of errors. Errors have been associated with the following: the personality of the rater, the scale itself (ambiguity), the nature of the action or trait itself, and the opportunity afforded the rater for observation (Mehrens and Lehmann, 1973, pp. 356-358). Raters come to the rating tasks with personal bias that tend to influence ratings and cause errors. Some of these errors associated with the rater's personality go by the following names: the halo effect, severity effect, central tendency error, and logical error.[1]

---

[1]Various authors cover these common errors related to raters. For the reader unfamiliar with these errors the following sources may be consulted: Guilford, 1954, pp. 278-280; Isaac, 1971, p. 58; Mehrens and Lehmann, 1973, pp. 357-358; and Remmers, 1963, pp. 372-373.

The scale itself may be ambiguous and this has also been cited as a source of errors. Mehrens and Lehmann (1973) defined ambiguity as "wording and meaning of the traits being measured - such that the rater may be uncertain as to what it is he is really being asked to rate" (p. 356). Ambiguity may be in the trait or the frame of reference between different levels of the trait being rated. For example, the trait of aggressiveness may be ambiguous if left operationally undefined because aggressiveness can be viewed as a positive trait (appropriately self-assertive) and a negative trait (hostility). An example of the frame of reference ambiguity could relate to the criteria of superior, good, and inferior as they have been related to swimming. What each of these terms meant in relation to swimming depended upon both the nature of the swimming trait being measured (type of stroke - crawl versus dog paddle) and the age and ability of the rater (60 year old man with a heart condition versus 20 year old man with no heart condition).

The nature of the trait or action being rated may cause errors. It has been well documented that as a trait moves from a unitary action to a series of actions the complexity of the rating and the possibility of errors increases.

The length of observation also has a definite bearing on the amount of error that may enter into ratings.

This is

observa

observa

ject be

Wears

V

Gulfo

pp. 35

to be

and fe

of rat

tion

This is especially true when ratings have been based on observation alone without either a chance for extensive observation and/or a chance to ask questions of the subject being observed (Mehrens and Lehmann, 1973, p. 358).

Means to Overcome These Errors

Various authors (Borg and Gall, 1971, pp. 234-239; Guilford, 1954, pp. 271-298; Mehrens and Lehmann, 1973, pp. 357-361) agreed that careful definition of behaviors to be rated and training of raters were the most effective and feasible means of improving the reliability and validity of ratings. Some points to consider in the clear definition of traits or actions were the following:

- Identify the action to be rated as a simple, unitary aspects of behavior when possible. Horrocks and Schoonover (1968, pp. 447-448) found, the more complex the behavior, the more difficult it was to operationally define and obtain high inter-rater reliability.
- Define the action or trait to be rated as much as possible in operational terms.
- Use common expressions and avoid technical jargon when possible for clarity.
- State the action to be rated as a question rather than a declarative statement.
- Establish the same set in the minds of raters by rating the most specific action possible and being the most descriptive of the response options (nature of correct and incorrect responses in the case of checklists).

- Stay clear of terms implying ethical, moral, or social evaluations as correct or incorrect responses, unless dealing specifically with such behaviors.

- Establish good cutting points for categories along an action or trait continuum (correct and incorrect responses in the case of a checklist). A good cutting point applies to a point or very short range on the continuum being rated.

Some points to consider in improving ratings by means of a training session were the following:

- Discuss the checklist form with the raters, describing each item sufficiently to develop a thorough understanding of why the data is being collected, under what conditions the observation should take place, what is to be observed, and how it is to be recorded.

- Comment frequently during the training session on the value of accurate and honest reporting.

- Discuss the kinds of errors usually committed by raters either through personal bias and/or logical errors and how they might be avoided, or at least minimized.

- Conduct some "dry runs" to give raters practice. If possible videotape testing sessions to serve as material for "dry runs."

- Provide discussion after "dry runs" to further clarify rating practices and procedures.

- Calculate inter-rater reliabilities after practice sessions to estimate the degree to which observers are developing a common frame of reference.

che

bil

app

rem

The above enumerated recommendations for better written checklists and adequate training sessions offer the possibility for improved consistency of ratings. When they are applied in a professional manner, rater agreement should remain consistently high across subjects rated.

SUMMARY

The review of literature covered two seemingly unrelated topics - historical changes in early childhood education's mission and the error associated with checklists. The linkage between the topics was shown to be due to the compensatory educational nature of present-day preschool programs that require screening (sometimes by means of checklists) of potential participants. Early childhood education as an organized enterprise in the United States has a short history. The first such endeavors started in the 1920's as nursery schools attached to universities and colleges for research purposes. Their curriculum consisted of habit training and the promotion of physical health. Federal funding during the depression years and then again during World War II increased the number of nursery schools but changed little their curriculum other than an increase of an emphasis on socio-emotional growth in the 1940's. During the 1960's, a combination of academic and socio-political forces finally lead to massive federal funding of compensatory education preschool programs to prevent school failure among lower-lower-class children. The curriculum of these preschools emphasized cognitive growth and other abilities necessary for success in schools while drawing curriculum elements from the former nursery school movement. Present-day preschools enjoy a great

variety in the composition of curriculum, but for the most part focus upon the following four areas of development: socio-emotional, perceptual-motor, cognitive, and language. Any readiness screening device for preschool education should be broad enough to cover these areas of child development.

Further such screening may be accomplished by means of a checklist as a means to measure readiness skills. While checklists at first glance appeared to be a simple measurement tool, errors of measurement were possible from a number of sources. The personality of the rater, the scale itself, the nature of the action (or trait), and the opportunity afforded the rater were all classified as common sources of error. Specific suggestions from the literature were offered to combat these errors under the broad headings of clearer definition of action rated and provisions for training raters.

ard

Pre

cha

set

str

sum

to

eff

ing

str

int

ing

the

of

apr

CHAPTER 3: METHODOLOGY

This chapter provides a description of the methods and procedures used to further develop and evaluate the Prekindergarten Readiness Screening Device (PRSD). The chapter is divided into the following major sections: setting, instrumentation, rater reliability study, construct validity study, predictive validity study, and summary. The three study sections provide the details to the three major thrusts of this developmental research effort: 1) determine rater reliability by raters viewing a common videotaped testing session, 2) estimate construct validity by examining percent passing by age interval, and 3) assess predictive validity by correlating two achievement instruments. Each study section has discussion under the following headings: test evaluation questions, subjects, procedures, and analysis. A number of measurement instruments are referred to by their acronyms in the following text.[1]

---

[1]These instruments and their associated acronyms follow: the Metropolitan Achievement Test (MAT), the Prekindergarten Readiness Screening Device (PRSD), and the Prekindergarten Saginaw Objective Referenced Test (PSORT).

SETTING

The study took place in Saginaw, Michigan, a mid-western industrial city of approximately 85,000. During the past decade increasing numbers of middle class white and middle class black families have been moving out from the city school district into the surrounding suburbs. Unskilled and semiskilled workers who have found employment in one of the city's three foundries have moved their families into the vacant homes left by the flight to the suburbs.

The study's data came from the School District of the City of Saginaw's Title I Prekindergarten program. This program, which has been in operation for the past twelve years, was funded through Title I of the Elementary and Secondary Education Act. It was designed to provide four year olds with an environment that would enable them to be ready for entry into school. The ultimate purpose of this program was to prepare inner city children, many of whom came from backgrounds that may not have equipped them with the skills necessary for success in school, for entry into kindergarten. After a year of prekindergarten, it was hoped these children would at least be on a par with other five year olds as they entered kindergarten.

Since September, 1977, the Saginaw Prekindergarten Program has been designated by the Michigan State Department of Education as a Demonstration Site in the Michigan Adoption Program. In order to achieve this distinction, it had to meet the Department's criteria for an exemplary program, including a comprehensive validation of its outcomes. What this meant was that the Prekindergarten Program had proven so successful over the past several years that it was now recognized as a model program that could be adopted by local districts around the State of Michigan.

At the time of the study there were approximately 420 children enrolled at the program's twelve sites: Baillie, Coulter, Emerson, Haley, Houghton, Jones, Longfellow, Loomis, Morley, Potter, Rouse and Salina (see Appendix D for the Saginaw day school elementary official membership by grade and racial and ethnic counts by building).

The Prekindergarten staff included a supervisor, certified early childhood teachers, teacher aides, a graphic arts aide and a secretary. With the exception of two half-day sites, each school operated two sessions, one from 9:00 a.m. to 11:30 a.m. and another from 12:30 p.m. to 3:10 p.m. Class sizes ranged up to 20 children per half-day session.

The program's instructional approach closely followed many aspects of Piagetian Theory. Jean Piaget, one of the world's foremost child psychologists, spent years studying the development of young children. He found that all children pass through certain stages of intellectual development. Growth, according to Piaget, was a process of exploring, manipulating, adapting, and assimilating the environment, and children go through various stages in an attempt to integrate or organize what they have observed.

The rationale for this approach was explained by Almy (1966, p. 127).

> Piaget's Theory clearly carries the implication that the young child has as much, if not more, to learn from his own active encounters with his physical environment and from his exchanges with his peers as he has from the adult. But adults, parents, and teachers ... are constantly responsible for decisions that determine the nature of the child's encounters and exchanges.

The idea was to make children feel good about themselves and about going to school. Teachers encouraged the children to experience as many activities and aspects of the program as possible without fear of failure. Activities were centered around three domains: psychomotor (small muscle and large muscle body movement), affective (social, attitudinal and emotional), and cognitive (reasoning and knowledge). (Appendix E presents the Title I Prekindergarten Activity Observation Checklist which gives

the broad types of activities in the three domains and gives a key for categorizing specific activities into these areas.) Thus in prekindergarten the children were in essence, learning to learn.

Language development was an important ingredient in the program and teachers were encouraged to stress it daily. Many children came from environments that place little emphasis on verbal communication. As a result, they were often unable to express themselves well verbally. One of the major goals of the program was to surround the children with language while they were in the classroom in an attempt to increase both their vocabulary and self-confidence.

Another major goal was parent involvement. Each teacher was required to make at least one home visit to meet parents and to get an idea of each child's home situation. Parents were encouraged to come to class with their children and participate. All prekindergarten parents were urged to take an active role in the child's education. Materials were sent home regularly so that parents and children could work together on various activities. The parent program was designed to teach parents to work more effectively with their children.

The program had goals and objectives (see Appendix F) that children worked on throughout the year. The structure of the program was in its planning and delivery system.

The above described setting should provide background into specifically what type of community Saginaw is and also the components of its operational prekindergarten program. However, the major purpose of this chapter is to provide a description of the methods and procedures employed in the three studies of the PRSD. Before a detailed description of the three studies is offered, an explanation of the instrumentation is furnished.

INSTRUMENTATION

This section focuses on the three measurement instruments employed in the studies to be described in the next three sections. First, the theoretical and historical background related to the development of the PRSD will be discussed. Since the further development and evaluation of the PRSD is the aim of this study, the instrument itself is described in some detail. Next both of the criterion measures for the predictive validity portion of the study are described and reliability estimates are presented.

Prekindergarten Readiness Screening Device (PRSD)

PRSD (see Appendix C for copy) is an individually administered 27 item applied performance checklist with standardized spoken directions and a statement of an acceptable response for each item. The instrument was designed to assess the entry behavior in early September of prekindergarten aged pupils who are potentially eligible for participation into Title I prekindergarten programs.

The skills included on the PRSD are those that research findings and teacher observations have determined that most four year olds possess (see Appendix G for research references). A small percentage of three and five year old skills were also included to give a broader range to the behaviors assessed. Table 3.1 which follows gives the behaviors requested of the pupils for each question.

TABLE 3.1. ITEM NUMBERS AND SPECIFIC BEHAVIORS REQUESTED OF PUPILS ON THE PREKINDERGARTEN READINESS SCREENING DEVICE (PRSD).

Item Number	Behavior Requested
1	Say last name
2	Say age
3	Point to neck
4	Identify body part
5	Tell function of body part
6	Pick up same colored object
7	Pick up object colored same as named color
8	Say color of object
9	Say color of object
10	Count to five
11	Give me four blocks
12	Say number of remaining blocks
13	Pick up blocks
14	Walk backwards
15	Carries out 3-part command
16	Tell what books are for
17	Draw cross given model
18	Hop after demonstrated
19	Throw ball five feet
20	Point to shape like a wheel
21	Point to shape like a tent
22	Point to shape like a stick
23	Tell which of two objects is bigger
24	Draw diamond given model
25	Tell nursery rhyme, song, poem
26	Say yellow jello
27	Verbal response to the above questions given in more than just one word responses

The PRSD resulted from the pilot form of the instrument developed jointly in 1975 by the Saginaw prekindergarten staff and the present author. The pilot version underwent further revisions in 1976 based on teacher observations gained in the prior administration. In June of 1977, an indepth review was undertaken by the author of the test by means of a training videotape developed to aid in the administration of the instrument. From discrepancies noted between the instrument and the videotape it was decided that administration and scoring directions needed to be further standardized if the PRSD was to yield consistent results. Standardized directions and scoring instructions were developed prior to September, 1977 and have been used since.

Prekindergarten Saginaw Objective Referenced Test (PSORT)

PSORT (see Appendix H for a copy) is an objective referenced test of 31 items dealing with both psychomotor and cognitive outcomes in prekindergarten program areas. The first 17 items measured the program's nine cognitive product objectives while the remaining 14 dealt with the four fine and gross motor objectives (see Appendix F for a copy of the product objectives and summary of mastery criteria). PSORT was an applied performance checklist with standardized directions and a statement of an acceptable response for each item much like the PRSD. The

major difference between the PSORT and PRSD is that PSORT was designed as an outcome measure sensitive to changes brought about by Saginaw's Title I Prekindergarten Program within a narrow range of behavior deemed appropriate to the program. The PRSD sampled a wider range of behaviors and was not specifically designed to be sensitive to a particular educational program.

Reliabilities of PSORT ratings were determined by having 20 raters view a videotape of two students and rate these students on the basis of PSORT scoring directions. Intra-class correlations as described by Ebel (1951) were calculated with the reliability of ratings being .761. and the reliability of average ratings being .984.

Metropolitan Achievement Tests (MAT)

The MAT used in this study is the preprimer level survey battery of the 1978 edition of the Metropolitan Achievement Tests published by the Psychological Corporation. The MAT was employed as part of the official testing program of the Saginaw Public Schools. A group of teachers representative of Saginaw teachers had chosen MAT because its items best matched Saginaw's curriculum in kindergarten and first grades. The MAT served as a criterion measure for the predictive validity of the PRSD. The preprimer level test covered important skills and content areas normally taught at the beginning of

kindergarten through the middle of kindergarten (Prescott, Balow, Hogan, and Farr, 1978). This test was designed to yield both norm-referenced and criterion-referenced information like many of the newer achievement tests.

A study of the reliability of the MAT's total test score was undertaken with the 346 potential prekindergarten students who were first screened in September, 1978 on the PRSD. The Kuder-Richardson 20 estimated reliability for the April, 1980 administration of MAT on the former screened students was .909. The 346 pupils over all scored at the 22 percentile on the MAT.

The next section provides a description of the rater reliability study. This study is the first of three studies to be described in the process of the further development and evaluation of the PRSD.

RELIABILITY

The estimation of rater reliability was one of three studies undertaken to further develop and evaluate the PRSD. This study consisted of teachers and aides rating the performance of two children after viewing their videotaped testing sessions. A short training period was held prior to scoring the two children on videotape. Intraclass correlations were calculated. Further details of the study follow under the headings of test evaluation question, subjects, procedures, and analysis.

Test Evaluation Question

This study was undertaken to answer the following question:

> What is the estimated rater reliability of the PRSD?

This question was formulated on the basis of the commonly held assumption that an applied performance test with standardized administration and scoring directions can bring about consistent results with trained staff. However, a check of the level of reliability of measurement seemed necessary because both standardization and training are relative to the context of the measurement and the population being measured. Consistent results were defined as reliability coefficients to be in the upper range of the r values, usually .70 to .98 (Guilford, 1973, p. 92).

Subjects

Nineteen adults (eleven teachers and eight aides) employed by the Title I Prekindergarten Program comprised the sample for the rater reliability portion of the study. All raters were female. All eleven teachers (100%) and four aides (50%) had administered the PRSD prior to the rater reliability study. Approval for participation of subjects was obtained from both the School District of the City of Saginaw and the University Committee on Research Involving Human Subjects (see Appendix I for copies of both letters).

Procedures

Teachers and aides participated in a rater reliability study on the morning of November 21, 1980. A training session of approximately fifteen minutes preceeded the actual study. During the training session the following topics were covered: purpose and rationale of the instrument, use of answer sheet to indicate responses, correct responses for the items, and what can be done to avoid typical rating errors. After the training session, all participants scored two videotaped testing sessions using answer sheets. These adult subjects were asked to score the PRSD and PSORT on the basis of commonly viewed videotaped testing sessions of two pupils each for both instruments. This participation was part of a training exercise that was scheduled on the part of the

prekindergarten supervisor. After the answer sheets were collected, a short discussion was held at the end of each PRSD screening session to obtain ways to improve the scoring and/or administration directions.

Analysis

The rater reliability was estimated by means of the intra-class correlation coefficient for both an individual rater and the average of the group. The rater reliability has been the critical estimate of reliability for checklists due to errors related to rater misinterpretation of checklist items.

The assumptions underlying the use of intra-class correlation for reliability estimates are the following: the error of measurement is uncorrelated with the true score, the sample of people on whom the observations are made is a random sample from a population of people to which inferences are to be made, the sample of raters used is a random sample from a population of comparable raters, and that the within-person variance may be pooled to provide an estimate of its magnitude (Winer, 1971, p. 286).

In addition, the percentage of raters scoring each item consistent with the original rater (the prekindergarten supervisor) over the two test administrations were reviewed to identify questions that seemed to be particularly troublesome to obtain scorer agreement. These items then

will be compared with error prone items mentioned in the feedback session after each scoring session.

CONSTRUCT VALIDITY

The second study undertaken was the exploration of one aspect of the PRSD's construct validity. This study involved testing if a monotonically increasing relationship existed between percent passing an item and each progressively higher chronological age interval. A test for trend in ordered contingency tables was used to test for significance of the monotonically increasing relation between proportion passing and age. Further details of the study follow under the headings of test evaluation question, subjects, procedures, and analysis.

Test Evaluation Question

The construct validity study sought to test the following hypothesis stated in narrative as well as symbolic forms.

Hypothesis 1.

> Null Hypothesis: No monotonically increasing relationship is shown in the proportion of pre-kindergarten subjects passing an item on the PRSD at each progressively higher chronological age interval.
>
> Symbolically: $H_o : P_1 \geq P_2 \geq P_3 \geq P_4$
>
> Legend: P_k = proportion of pupils passing an item at a given chronological age interval k (where k = 1 represents the youngest age interval)

Alternate Hypothesis: A monotonically increasing relationship is shown in the proportion of pre-kindergarten subjects passing an item on the PRSD at each progressively higher chronological age interval.

Symbolically: $H_1 : P_1 < P_2 < P_3 < P_4$

Legend: P_k = proportion of pupils passing an item at a given chronological age interval k (where k = 1 represents the youngest age interval)

Hypothesis 1 was based on the common practice for preschool and infant tests to be validated in terms of age differentiation (Anastasi, 1970, p. 474). In other words, test results at the item level were checked against chronological age to determine whether performance showed an improvement with advancing age. However, age progression of items is a necessary but not sufficient condition for construct validity. Thus, if the percentage passing each item fails to improve with age, such a finding indicates that the test is not a valid measure of readiness since it is supposed to increase with age with a broad sampling of individuals. On the other hand, to prove that a test measures something that increases with age does not define the particular area covered by the test very precisely. But it may be looked upon as providing one piece of evidence related to showing that the test has one element related to the nomological network related to the construct of readiness.

Subjects

A total of 1,415 potential Title I preschool pupils served as the sample for estimating the construct validity of the PRSD. These subjects were screened for possible participation during the three consecutive school years of 1978-79, 1979-80, and 1980-81.

Almost all subjects screened were four years old by December 1 and would not be five years old by December 2 in the year of their screening. Approximately 50% were females (see Appendix J for exact counts by sex and age). Racial ethnic background of the subjects was approximately the following: 77% Black, 16% Latino or Spanish, 6% White, 0.2% American Indian, 0.1% Asian, and 0.6% unknown (see Appendix K for exact counts). Almost all the children tested resided in the Title I attendance area of the School District of the City of Saginaw. The Title I attendance area has been so designated because of the low academic achievement results of the elementary student populations. The large majority of low social economic status inhabitants of Saginaw live in these attendance areas (see Appendix L for the latest available census data for this area).

Approval for participation of subjects was obtained from both the School District of the City of Saginaw and the University Committee on Research Involving Human Subjects (see Appendix I for copies of both letters).

Procedures

Potential prekindergarten pupils were screened during September and October for three consecutive years (1978, 1979, and 1980) using the PRSD. Teachers and aides screened the pupils as part of the selection process for participation into the Title I Prekindergarten program. The test administrators were given a training session on how to administer the PRSD in early September, prior to the starting of the screening process each year. In addition, the researcher and the program supervisor observed testing sessions and afterwards corrected administration and scoring errors observed for each of the three years. The resulting answers along with demographic information of the pupil (month and year of birth, sex, and race) were recorded directly on a standard optical scanning answer sheet. These answer sheets were machine scored locally.

Analysis

Hypothesis 1 related to each item was tested by means of a chi-square test followed by a test for trend in an ordered contingency table (Marascuilo and McSweeney, 1977, pp. 198-202). This test was chosen because it is the only easily available statistical test to test the particular hypothesis of interest.

Specifically, the hypothesis was tested in the following manner: The observed frequencies were displayed in a

contingency table constructed of chronological age groups (4 years or less, 4 years 1 month to 4 years 3 months, 4 years 4 months to 4 years 6 months, and 4 years 7 months or more; age as of September of each school year) across the top crossed by pass and fail down the side for each of the PRSD items. Figure 3.1 below presents the contingency table layout used for each item.

Figure 3.1. Chi-Square Contingency Table Layout Used for Each Item of the PRSD.

Answer Status On PRSD Item	Age Group (Year-Month)				Total
	4-0 or less	4-1 to 4-3	4-4 to 4-6	4-7 or more	
Pass					
Fail					
Total					

Chi-square values were calculated on the basis of expected frequencies determined from marginal totals from each contingency table. By using linear regression theory the portion of the chi-square attributable to the monotonicity of the relationship was determined by calculating $\hat{B}$ and $SE_{\hat{B}}$. The resulting value of χ^2 due to the monotonicity of the linear trend analysis was tested at alpha = .05. Cramer's contingency coefficients were calculated for the variation

due to the monotonicity of the regression for each test item. If the majority of Cramer's contingency coefficients were .20 or greater then separate percentiles by age group would be calculated. If the majority of Cramer's contingency coefficients were below .20, then a Pearson product moment correlation coefficient would be calculated for the association of chronological age (age expressed as a decimal, for example, 4 years 5 months = 4.05) and total PRSD raw score. If this correlation was .20 or greater then separate percentiles would still be calculated for each age group, on the other hand, if the correlation was less than .20 then only percentiles for the total group would be calculated.

Kuder-Richardson 20 and split-half (even versus odd) reliabilities were calculated using the data set from the construct validity study. The obtained split-half reliability was corrected by means of a derived formula from the Spearman-Brown prophecy formula by Horst (1951) which estimates total test reliability from parts of unequal length (see Appendix M for formula).

PREDICTIVE VALIDITY

The final study undertaken was the assessment of the PRSD's predictive validity. The results of the PSORT administration of October, 1978 and April, 1979 plus the MAT administration of April, 1980 served as the achievement criterion measures. The Pearson product moment correlation coefficients were calculated for each of the criterions. Further details of the study follow under the headings of test evaluation questions, subjects, procedures, and analysis.

Test Evaluation Questions

The predictive validity study sought to test the following two hypotheses stated in narrative as well as symbolic forms.

Hypothesis 2.

Null Hypothesis: The correlation coefficient between the PRSD and the pre-test PSORT will not be greater than zero.

Symbolically: $H_o : \rho_{12} = 0$

Legend: ρ = population correlation coefficient
1 = PRSD
2 = pre-test PSORT

Alternate Hypothesis: The correlation coefficient between the PRSD and the pre-test PSORT will be greater than zero.

Symbolically: $H_2 : \rho_{12} > 0$

Legend: ρ = population correlation coefficient
1 = PRSD
2 = pre-test PSORT

Hypothesis 3.

Null Hypothesis: The correlation coefficient between the PRSD and the post-test PSORT will not be greater than zero.

Symbolically: $H_0 : \rho_{13} = 0$

Legend: ρ = population correlation coefficient
1 = PRSD
3 = post-test PSORT

Alternate Hypothesis: The correlation coefficient between the PRSD and the post-test PSORT will be greater than zero.

Symbolically: $H_3 : \rho_{13} > 0$

Legend: ρ = population correlation coefficient
1 = PRSD
3 = post-test PSORT

Hypothesis 4.

Null Hypothesis: The correlation coefficient between the PRSD and the MAT will not be greater than zero.

Symbolically: $H_0 : \rho_{14} = 0$

Legend: ρ = population correlation coefficient
1 = PRSD
4 = MAT

Alternate Hypothesis: The correlation coefficient between the PRSD and the MAT will be greater than zero.

Symbolically: $H_4 : \rho_{14} > 0$

Legend: ρ = population correlation coefficient
1 = PRSD
4 = MAT

Hypotheses 2, 3, and 4 were based on the assumption that school readiness are necessary to show success on later measures of school achievement.

Subjects

A total of 396 Title I preschool pupils served as the sample for assessing the predictive validity of the PRSD. These subjects were screened for possible program participation during the 1978-79 school year.

Almost all subjects screened were four years old by December 1 and would not be five years old by December 2 in the year of their screening. Approximately 50% were females. See Appendix J for exact counts by sex and age. Racial ethnic background was highly minority with approximately 5% being caucasion or white (see Appendix K). Almost all the children resided in Title I attendance area of the School District of the City of Saginaw where the large majority of low social economic status inhabitants of Saginaw live (see Appendix L for the latest available census data of this area).

Approval for participation of subjects was obtained from both the School District of the City of Saginaw and the University Committee on Research Involving Human Subjects (see Appendix I for copies of both letters).

Procedures

The predictive validity study involved testing pupils with the PRSD and then testing them with the PSORT and MAT as criterion measures. The pupils were screened during September and October, 1978 using the PRSD. The

prekindergarten staff members responsible for the screening attended a training session on the administration and scoring of the PRSD. The teachers were observed screening and any errors in administration were corrected. The results of the screening sessions were coded on a standard optical scannable answer sheet. These answer sheets were machine scored locally.

The same pupils who remained in the prekindergarten program were then tested using the PSORT during October, 1978 (pre-test) and April, 1979 (post-test) as part of the Title I program evaluation activities. Prekindergarten staff members received training on how to administer the PSORT prior to pre-testing. The test administrators recorded the correctness of responses directly on answer sheets for both the pre- and post-tests. These answer sheets were machine scored locally for both test administrations.

The pupils still remaining in the Saginaw Public Schools were tested during April, 1980 using the MAT as part of the scheduled testing program for kindergarteners. All kindergarten classroom teachers attended an hour long training session on how to administer the MAT which highlighted the use of the test manual as a means to standardized the testing session. Pupils marked directly on the answer booklet which were machine scored by the test publisher.

Analysis

Hypotheses 2, 3, and 4 were tested in the following manner: Pearson product-moment correlations were computed between PRSD raw scores and the raw scores of the pre-test PSORT, post-test PSORT, and MAT. The Fisher r to Z transformation, which assumes the (X, Y) joint events have a bivariate normal distribution in the population, was used to test for significance. The alpha level of significance chosen was .05.

SUMMARY

This chapter provided a description of the methods and procedures used to further develop and evaluate the Prekindergarten Readiness Screening Device (PRSD). The study took place in Saginaw, Michigan, a mid-western industrial city of approximately 85,000. The School District of the City of Saginaw's Title I Prekindergarten program served as the setting for the study. Both prekindergarten staff members and potential pupils served as subjects. Data from the 1978-79, 1979-80, and 1980-81 served as the basis for the three major developmental thrusts of the study - 1) determine rater reliability, 2) estimate construct validity, and 3) assess predictive validity. The PRSD, the Prekindergarten Saginaw Objective Referenced Test (PSORT) and Metropolitan Achievement Tests (MAT) were described in some detail with reliability estimates for Saginaw Title I pupils of .761 and .909 offered for PSORT and MAT respectively.

The rater reliability study consisted of raters scoring two videotaped testing sessions. A short training session was held prior to scoring two children on videotape. Intra-class correlations were calculated. Also percent agreement of these raters were compared to the original rater and items identified as difficult to score during the scoring sessions.

secut

const

the c

trend

inter

cells

level

Septe

serve

the c

Octob

the 7

trat

Meas

cien

test

ferey

form

The total set of the PRSD results from the three consecutive school years mentioned above was employed in the construct validity study. This data set was used to study the construct validity of the items by means of a test for trend in ordered contingency tables. Chronological age intervals crossed with pass and fail on PRSD formed the cells for the ordered contingency table for each item. The level of significance was $\alpha = .05$.

In the predictive validity study the results of September and October, 1978 administration of the PRSD served as the test scores (or the X variable) to predict the criterion. The results of the PSORT administration of October, 1978 (pre-test) and April, 1979 (post-test) for the Title I Prekindergarten program plus the MAT administration of April, 1980 served as the achievement criterion measures. The Pearson product-moment correlation coefficients were calculated for each of the criterions. To test whether these correlations were significantly different than a zero correlation, a Fisher r to Z transformation was used. The level of significance was $\alpha = .05$.

CHAPTER 4: ANALYSIS OF RESULTS

In this chapter, the results of the developmental and evaluative methods employed on Prekindergarten Readiness Screening Device (PRSD) data are presented and discussed. The chapter is divided into the following sections: reliability, construct validity, and predictive validity. These three sections are offered to shed some light upon the quality of measurement offered by the PRSD. The measurement instruments used in the study are referred to by their acronyms in the following text.[1]

---

[1]These instruments and their associated acronyms follow: the Metropolitan Achievement Test (MAT), the Prekindergarten Readiness Screening Device (PRSD), and the Prekindergarten Saginaw Objective Referenced Test (PSORT).

RELIABILITY

Reliability estimates for psychological and educational tests can be thought of as a statistical index of measurement quality. While reliability is a necessary but not sufficient condition to judge the overall quality of a test, it is one estimate of quality most test constructors must consider early in the development process.

The reliability estimate most appropriate for a particular test depends on the type of test and its administration and scoring environment. Performance checklists, like the PRSD, seem most prone to errors by raters in the process of evaluating performance. The group of raters must have a common frame of reference relevant to scoring behaviors in response to checklist items.

The intra-class correlation provided an estimate of rater reliability for both a single rater and a group of raters by an analysis of variance technique. This seems much more efficient and accurate technique than calculating Pearson correlations between each possible pair of raters.

In the present study, 19 raters (teachers and paraprofessionals) rated the behavior of two children by viewing a videotaped testing session of each.

Table 4.1 shows the variances used to calculate the intra-class correlations for a single rater and the average for the group of raters.

TABLE 4.1. VARIANCES USED TO ESTIMATE INTRA-CLASS RELIABILITY OF PRSD.

Source	Sum of Squares	Degrees of Freedom	Variance
From children	290.13	1	290.13
From raters	53.05	18	2.947
Remainder	19.37	18	1.076
Total	362.55	37	*

*Variance not needed and not computed.

The formula and calculation for the intra-class coefficient for a single rater follows:

$$\bar{r}_{11} = \frac{V_c - V_e}{V_c + (k-1)\,Ve} = \frac{290.13 - (1.076 + 2.947)}{290.13 + (19-1)(1.076 + 2.947)} = .789$$

where

$\bar{r}_{11}$ = reliability of ratings for a single rater

V_c = variance for children

Ve = variance for raters + variance for remainder = variance for error[1]

k = number of raters

[1]Even though the ratings for the rater reliability study were collected differently than typically employed (each rater rated each child rather than each rater rating only a unique set of children), still it is appropriate to calculate the error variance by including the rater variance along with the remainder variance. This inclusion is necessary since differences from rater to rater in general level of rating do lead to corresponding differences in selection, the between raters variance should be included in the error term.

The formula and calculation for the intra-class coefficient for the means of the two ratings for each child follows:

$$r_{kk} = \frac{V_c - Ve}{V_c} = \frac{290.13 - (1.076 + 2.947)}{290.13} = .973$$

where

r_{kk} = reliability for mean ratings from k raters

V_c = variance for children

Ve = variance for raters + variance for remainder = variance for error[1]

k = number of raters

Both obtained intra-class coefficients of .789 (for a single rater) and .973 (for an average of all the raters) were within the band defined previously as characteristic of consistent measurement. Guilford (1973, p. 92) stated that typically measurement procedures with reliability coefficients between .70 to .98 tend to yield consistent results. It is noteworthy that the obtained coefficients are within this band. The rater reliability for a single person (r = .789) is the best estimate of reliability when only one individual will screen each child as is done in Saginaw. The rater reliability of the mean of all raters (r = .973)

---

[1]Even though the ratings for the rater reliability study were collected differently than typically employed (each rater rated each child rather than each rater rating only a unique set of children), still it is appropriate to calculate the error variance by including the rater variance along with the remainder variance. This inclusion is necessary since differences from rater to rater in general level of rating do lead to corresponding differences in selection, the between raters variance should be included in the error term.

is the best estimate of reliability when the same child will be screened by two or more persons. Thus each intra-class coefficient is more valid in one situation than another.

The reader is reminded that other reliability estimates (the Kuder Richardson-20 and split half reliabilities) will be presented in the construct validity section of this chapter.

In addition, the percentage of raters agreeing with the original rater was used as a means to explore if any item tended toward fostering rater disagreement and thus inconsistency in the measurement procedure. The percentage was calculated by dividing the number of raters in agreement with the original rater by the total number of raters and multiplying this answer by 100. The comments made after the videotaped screening session were gathered on possible ways to improve the scoring and/or administrative consistency. Table 4.2 below summarizes this information on both pupils tested plus comments offered to making testing more uniform in the future.

TABLE 4.2. PERCENT AGREEMENT OF VIDEOTAPE RATERS TO ORIGINAL RATER (N = 19).

PRSD Item Number	Behavior Requested	Percent Agreement		Comments
		Pupil A	Pupil B	
1	Say age	94.7	68.4	
2	Say name	100.0	94.7	
3	Point to neck	100.0	100.0	
4	Identify body part	100.0	100.0	
5	Tell function of body part	94.7	100.0	
6	Pick up same color	100.0	100.0	
7	Pick up named color	94.7	100.0	
8	Say color of object	78.9	100.0	
9	Say color of object	100.0	100.0	
10	Count to five	100.0	84.2	
11	Give me four blocks	94.7	100.0	
12	Say number of blocks	94.7	100.0	
13	Pick up blocks	94.7	100.0	
14	Walk backwards	10.5	21.1	
15	Carry out 3-part command	94.7	84.2	
16	Tell what books are for	94.7	100.0	
17	Draw cross given model	89.5	42.1	
18	Hop after demonstrated	100.0	100.0	Is hopping in place correct or is a forward hop only acceptable?
19	Throw ball five feet	100.0	100.0	Should you score first throw or allow multiple throws?
20	Point to shape like a wheel	100.0	100.0	
21	Point to shape like a tent	100.0	84.2	
22	Point to shape like a stick	100.0	100.0	
23	Tell which is bigger	94.7	100.0	
24	Draw diamond given model	52.6	94.7	Hard to demonstrate tracing around template.
25	Tell nursery rhyme, song, poem	10.5	100.0	Reword and simplify - Can you tell me a song or poem?
26	Say yellow jello	89.5	63.2	
27	More than one word response	5.3	94.7	

Percent agreements as shown in Table 4.2, ranged from 5.3% to 100.0% and 21.1% to 100.0% for pupils A and B respectively. If a standard of 75% is used for comparison purposes for adequate rater agreement, a total of 23 out of 27 items (85.2%) had percent agreements above the 75% standard for both pupils rated. Only item 14 (walk backwards) had both percent agreements below this standard (10.5 and 21.1 percent for pupils A and B respectively). The explanation of why item 14 tended toward such disagreement may come from one or both of two plausible explanations. The perspective of the videotape raters was somewhat different than the original rater rating the child's live performance thus their view of the children's performance could have been responsible for the difference. The other explanation comes from the research of Horrocks and Schoonover (1968, pp. 447-448) that the more complex the behavior, the more difficult it was to operationally define and obtain high inter-rater reliability. During the three years of observing PRSD screening sessions, there was evidence that the test administrators had trouble in demonstrating the behavior properly before asking the child to walk backwards. If the walking backwards was a difficult task for teachers, then a leniency error toward the children may have been evident in their rating behavior.

A comparison of comments relating to scoring or administration problems and items with percent agreements

below the 75% standard was also undertaken. Of the four questions receiving comments, only items 24 (draw a diamond given model) and 25 (tell nursery rhyme, song, or poem) also received low ratings (of 52.6 and 10.5 percent respectively). The comments of clearer instructions on how to demonstrate the tracing of the diamond shape and simplifying the directions for requesting the child to tell a song or poem may make it easier for children to complete the PRSD tasks. This in turn may make it easier for the raters to score the items since objective standards for an acceptable response are provided. Without these modifications, some test administrators may be prone to resort to leniency because of their inability to provide the proper directions to the children being screened.

Overall, the PRSD seems to provide consistent measurement with rater reliabilities of .789 for a single person and of .973 for an average over all raters. The single person estimated reliability is appropriate when only one rater will screen a child and the average overall raters' estimated reliability is appropriate when two or more persons will screen each child. The percent agreement with the original rater tentatively indicated a number of items where agreement was lacking. It was hypothesized that agreement was lacking because of one or both of the following explanations. The videotape quality in terms of perspective, audio volume, and clarity may have caused some

of the disagreement because the original rater scored the live testing. The other explanation was that the lack of agreement could have resulted from observer leniency error due to the perception that either administration directions or scoring criteria were less than adequate for the children being screened to perform the task adequately and/or given credit for their attempts given the situation. Rater reliability as an index of measurement quality seems adequate to bring about consistent measurement generally, with a few items showing disagreement being the exceptions rather than the rule.

CONSTRUCT VALIDITY

The construct validity study explored one aspect of the readiness construct, namely the age progression of the percent passing each PRSD item. In other words, test results at the item level were checked against chronological age to determine whether performance showed an improvement with advancing age. However, age progression of items is a necessary but not sufficient condition for proof of the readiness construct. The results are provided as one element related to the nomological network related to the total construct of readiness.

The screening data from the three school years of 1978 through 1981 were used to study the age progression of each item by means of a test for trend in ordered chi-square contingency tables. Chronological age intervals crossed with pass or fail on PRSD formed the cells for the ordered contingency table for each item.

A test for linear trend (or more accurately termed test of monotonicity of a relationship) was calculated for each PRSD item to determine if the null form of the hypothesis should be rejected. In plain language, the null hypothesis was that no increase was shown in the percent of pupils passing an item at each progressively higher age interval. Hypothesis one is restated below in both narrative and symbolic forms.

Hypothesis 1.

Null Hypothesis: No monotonically increasing relationship is shown by the proportion of prekindergarten subjects passing an item on the PRSD at each progressively higher chronological age interval.

Symbolically: $H_o : P_1 \geq P_2 \geq P_3 \geq P_4$

Legend: P_k = proportion of pupils passing an item at a given chronological age interval k (where k = 1 represents the youngest age interval)

Alternate Hypothesis: A monotonically increasing relationship is shown by the proportion of prekindergarten subjects passing an item on the PRSD at each progressively higher chronological age interval.

Symbolically: $H_1 : P_1 < P_2 < P_3 < P_4$

Legend: P_k = proportion of pupils passing an item at a given chronological age interval k (where k = 1 represents the youngest age interval)

Appendix N presents the number and percent of the age groups passing each PRSD item. These data served as the information used in the test of linear trend. Table 4.3 below presents the results of the trend analysis.

TABLE 4.3. CHI-SQUARE TABLE FOR THE TEST OF LINEAR TREND AND CRAMER'S CONTINGENCY COEFFICIENTS FOR EACH PRSD TEST ITEM (N = 1,404).

Sources of Variation	df	Item 1 χ^2	C**	Item 2 χ^2	C	Item 3 χ^2	C	Item 4 χ^2	C	Item 5 χ^2	C
Due to monotonicity of regression	1	5.78*	.064	31.31*	.149	26.02*	.136	30.09*	.146	35.62*	.159
Departure from monotonicity of regression	2	4.53	.057	2.34	.041	1.90	.037	3.70	.032	4.15	.054
Total	3	10.31*	.086	33.65*	.155	27.92*	.141	33.79*	.155	39.78*	.168

Sources of Variation	df	Item 6 χ^2	C	Item 7 χ^2	C	Item 8 χ^2	C	Item 9 χ^2	C	Item 10 χ^2	C
Due to monotonicity of regression	1	34.21*	.156	45.59*	.180	38.86*	.166	36.36*	.161	22.91*	.128
Departure from monotonicity of regression	2	3.39	.049	.34	.016	1.23	.030	4.86	.059	.75	.023
Total	3	37.60*	.164	45.93*	.181	40.09*	.169	41.22*	.171	23.66*	.130

Sources of Variation	df	Item 11 χ^2	C	Item 12 χ^2	C	Item 13 χ^2	C	Item 14 χ^2	C	Item 15 χ^2	C
Due to monotonicity of regression	1	32.06*	.151	28.40*	.142	6.46*	.068	13.91*	.100	18.99*	.116
Departure from monotonicity of regression	2	3.79	.052	.53	.019	1.16	.029	.49	.019	1.99	.038
Total	3	35.85*	.160	28.92*	.144	7.62	.074	14.40*	.101	20.98*	.122

*$p < .05$.
**The amount of association given by each source of variation under the Cramer coefficient is not additive into a total figure because of the square root operation involved in the calculation of the Cramer coefficient.

Table 4.3. (Cont'd.)

Sources of Variation	df	Item 16 χ^2	C	Item 17 χ^2	C	Item 18 χ^2	C	Item 19 χ^2	C	Item 20 χ^2	C
Due to monotonicity of regression	1	47.27*	.183	45.70*	.180	22.86*	.128	4.50	.057	7.35*	.072
Departure from monotonicity of regression	2	2.82	.045	1.64	.034	5.50	.063	.87	.025	.72	.023
Total	3	50.09*	.189	47.34*	.184	28.36*	.142	5.37	.062	8.07*	.076

Sources of Variation	df	Item 21 χ^2	C	Item 22 χ^2	C	Item 23 χ^2	C	Item 24 χ^2	C	Item 25 χ^2	C
Due to monotonicity of regression	1	17.58*	.112	18.22*	.114	35.78*	.160	62.10*	.210	11.08*	.089
Departure from monotonicity of regression	2	1.04	.027	2.05	.038	6.52*	.068	..45	.018	3.70	.051
Total	3	18.62*	.115	20.27*	.120	42.30*	.174	62.55*	.211	14.78*	.103

Sources of Variation	df	Item 26 χ^2	C	Item 27 χ^2	C
Due to monotonicity of regression	1	22.18*	.126	1.45	.032
Departure from monotonicity of regression	2	2.59	.043	3.92	.053
Total	3	24.77*	.133	5.37	.062

*$p < .05$.

A quick glance at Table 4.3 reveals that the same type of data was provided for each of the 27 items of the PRSD. It may be instructive if an explanation of how each tabled number for a particular item was calculated using the actual data as an example.

Let item two be the example. The chi-square total value of 33.65 came from calculating the chi-square for the contingency table of item two formed by crossing the pass or fail status with the four age intervals. The contingency table of observed frequencies and marginal totals for item two is presented in Figure 4.1 below.

Figure 4.1. Observed Frequencies for the Contingency Table of Item Two of the PRSD.

Answer Status Y_2	Age Group (Year-Month) Y_1				Total
	4-0 or less	4-1 to 4-3	4-4 to 4-6	4-7 or more	
Pass	202	216	234	251	903
Fail	153	155	110	83	501
Total	355	371	344	334	1,404

The chi-square test also required the calculation of expected cell frequencies on the basis of the null hypothesis. The null hypothesis was that age was independent of the ability to pass an item, i.e., that the proportion of children

passing an item was the same regardless of whether they were 4 or less years, 4 years 1 month to 4 years 3 months, 4 years 4 months to 4 years 6 months, or 4 years 7 months or more old. To calculate the expected cell frequencies, the two marginal totals common to a cell of Figure 4.1 were multiplied by each other and then divided by the total number of children represented in the table. The contingency table of expected frequencies is presented in Figure 4.2 below.

Figure 4.2. Expected Frequencies for the Contingency Table of Item Two of the PRSD.

Answer Status	Age Group (Year-Month)				Total
	4-0 or less	4-1 to 4-3	4-4 to 4-6	4-7 or more	
Pass	228.32	238.61	221.25	214.82	903
Fail	126.68	132.39	122.75	119.18	501
Total	355	371	344	334	1,404

The degrees of freedom for an r x k contingency table such as Figures 4.1 or 4.2 was found by:

$$df = (r-1)(k-1) = (2-1)(4-1) = 3$$

where

r = number of rows (or classifications)
k = number of columns (or groups)

The computation of chi-square for the data in Figures 4.1 and 4.2 is straightforward:

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{k} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

$$= \frac{(202 - 228.32)^2}{228.32} + \frac{(234 - 238.61)^2}{238.61} + \ldots$$

$$+ \frac{(83 - 119.18)^2}{119.18} = 33.65$$

where

χ^2 = chi-square

O = observed cell frequency

E = estimated cell frequency

To determine the significance of chi-square = 33.65 when df = 3, a chi-square table was consulted to determine that when df = 3 and alpha = .05 that the observed value must be greater than chi-square = 7.82 to not accept the null hypothesis of no difference (indicated on Table 4.3 with an asterisk).

Since the hypothesis of interest involves an ordered relationship of age (Y_1) with respect to answering status on the PRSD (Y_2), a test for trend in a contingency table must be undertaken since the chi-square test of independence is insensitive to ordering of the columns or rows. To initiate the computations, the ordered classes of rows and columns were quantified by coefficients associated with the linear components from a standard table of orthogonal polynominals. According to the figures of Table A-9 (Marascuilo and McSweeney, 1978, p. 498), the linear coefficient for

K = 2 are -1 and 1, while for K = 4 the linear coefficients are -3, -1, 1, 3. With these scaled values and the frequencies of Figure 4.1 the following were calculated:

$$\Sigma Y_1 = 355(-3) + 371(-1) + 344(1) + 334(3) = -90$$

$$\Sigma Y_1^2 = 355(-3)^2 + 371(-1)^2 + 344(1)^2 + 334(3)^2 = 6,916$$

$$\Sigma Y_2 = 903(-1) + 501(1) = -402$$

$$\Sigma Y_2^2 = 903(-1)^2 + 501(1)^2 = 1,404$$

$$\Sigma Y_1 Y_2 = 202(-3)(-1) + 216(-1)(-1) + 234(1)(-1) + 251(-1)(3) + 153(-3)(1) + 155(-1)(1) + 110(1)(1) + 83(1)(3) = -420$$

With standardized linear regression theory, the slope of a regression line can be computed as:

$$\hat{B} = \frac{N(\Sigma Y_1 Y_2) - (\Sigma Y_1)(\Sigma Y_2)}{N(\Sigma Y_1^2) - (\Sigma Y_1)^2} = \frac{1,404(-420) - (-90)(-402)}{1,404(6,916) - (-90)^2} = -.064510$$

At this point, the nonparametric test departs from the classical procedure, in that:

$$S^2_{Y_2 \cdot Y_1} = \frac{N - 1}{N - 2}\left[S^2_{Y_2} - \hat{B}^2 S^2_{Y_1}\right]$$

is not used to compute $SE^2_{\hat{B}}$, since the determination of the partitioned chi-square component is made under the assumption that B = 0. When B = 0, it follows that $S^2_{Y_2 \cdot Y_1} = S^2_{Y_2}$, so that:

$$SE^2_{(\hat{B} = 0)} = \frac{S^2_{Y_2}}{(N - 1)\, S^2_{Y_1}}$$

For the observed data:

$$SY_1^2 = \frac{1,404(6,916) - (-90)^2}{1,404(1,403)} = 4.9253$$

$$SY_2^2 = \frac{1,404(1,404) - (-402)^2}{1,404(1,403)} = .9186$$

$$SE^2_{(\hat{B} = 0)} = \frac{.9186}{(1,404 - 1)(4.9253)} = .0001329$$

Under large-sample theory, $Z = \hat{B}/SE_{(\hat{B} = 0)}$ has a sampling distribution that is approximately N(0, 1), so that $\chi^2 = \hat{B}^2/SE^2_{(\hat{B} = 0)}$ is approximately chi-square with df = 1 degree of freedom. Thus the observed results:

$$\chi^2 = \frac{\hat{B}^2}{SE^2_{(\hat{B} = 0)}} = \frac{(-.0645)^2}{.0001329} = 31.31$$

To determine the significance of chi-square = 31.31 when df = 1, a chi-square table was consulted to determine that when df = 1 and alpha = .05 that the observed value must be greater than 3.84 to not accept the null hypothesis of no age progression related to PRSD item two (31.31 on Table 4.3 for item two shows this significance with an asterisk).

The variation due to departure from monotonicity of regression (age progression), as shown on Table 4.3 was calculated by simple subtraction.

chi-square non-age progression = chi-square total - chi-square age progression = 33.65 - 31.31 = 2.34

df chi-square non-age progression = df chi-square total - df chi-square age progression = 3 - 1 = 2

To determine the significance of chi-square = 2.34 when df = 2, a chi-square table was consulted to determine that when df = 2 and alpha = .05 that the observed value must be greater than 5.99 to not accept the null hypothesis (2.34 appeared on Table 4.3 for item two and lacked an asterisk, which indicated the departure from age progression was not significant).

Cramer's contingency coefficient (C) was also calculated for inclusion in Table 4.3. The general formula to determine this measure of association follows:

$$C = \sqrt{\frac{\chi^2}{N(L - 1)}}$$

where

L = the smaller of the number of rows or columns in the contingency table

N = total number of observations that make up the contingency table

χ^2 = chi-square

The calculations for Cramer's contingency coefficient for item two as shown in Table 4.3 follows:

$$C \text{ due to monotonicity} = \sqrt{\frac{31.31}{1,404(2 - 1)}} = .149$$

$$C \text{ due to departure from monotonicity} = \sqrt{\frac{2.34}{1,404(2 - 1)}} = .041$$

$$C \text{ total} = \sqrt{\frac{33.65}{1,404(2 - 1)}} = .155$$

Since the chi-squares of item two for variations due to monotoncity of regression (linear function of age progression)

and the total were found to be significant, the Cramer contingency coefficients of .149 and .155 calculated from the significant chi-squares are also significant by definition.

A review of all chi-square values for monotonicity in Table 4.3 indicates that the null hypothesis of no age progression related to the pupil's ability to pass a PRSD item failed to be accepted for 25 of the 27 PRSD items ($p < .05$). Only the results of item 19 (throw ball five feet) and 27 (more than one word responses) showed acceptance of the null hypothesis of no increasing ability to pass a PRSD item on the basis of age. The degree of association between age groups and the ability to pass PRSD items as shown by the Cramer coefficients ranged from .064 to .210 for the 25 items for which the alternate hypothesis was concluded. The average Cramer coefficients were .136 and .129 for the 25 items showing significance and all 27 items respectively. Only one of the items (item 24 - draw diamond given model) had a Cramer coefficient equal to or greater than .20.

The Cramer coefficients appear to be somewhat weak (average .129), however, it does seem there was some increase in the percentage of pupils passing PRSD as the next higher age groups were observed (as shown by the average proportion correct of .4046, .4492, .5136, and .5562 for each progressively higher age group over all items). It is true that the measure of dependence is small when compared with the possibility of total dependence, but it is large when compared with the possibility of no dependence.

There are at least two possible reasons, why the association shown by the Cramer coefficients appeared smaller than they could be under other circumstances. First, the Cramer coefficient is not directly comparable to the Pearson product moment correlation coefficient. Conover (1971, p. 177) notes that in general the Cramer coefficient has the desirable feature of being between 0 and 1.0 at all times like the Pearson coefficient, but it has the undesirable feature of depending on the number of rows and columns for its interpretation. The larger the number of rows and columns are, the larger chi-square in the equation for the Cramer coefficient tends to be, and division by (L - 1) only partially offsets this tendency. Thus judgments about size of the association cannot be tied directly to criteria used for judgments about Pearson product moment correlation coefficients. With a contingency table of 2 by 4 cells, like those used for the PRSD items, the size of the Cramer coefficient is limited to values less than one, which is unlike the Pearson product moment correlation coefficient.

Secondly, the age range used in the independent variable limited the variation in one or both of the variables. As a general rule, the smaller the range in both the independent and dependent variable, the smaller the numerical values of a measure of association, other things being equal. Appendix G, which shows norms reported by other researchers for items the same as, or similar to,

the PRSD items, shows most item norms spanning 6 to 12 or more months rather than the 3 month interval used in this study. It may be that the rate of development using 3 month intervals is too variable to define norms for any group of children more precisely and obtain strong measures of association. Overall, this research does tend to lend some support to the age progression of PRSD items.

Another portion of the construct validity study sought to use the Cramer coefficients as partial data to decide whether age group or total group percentile norms should be developed. The supposition was that each PRSD item alone might show enough association to support the age progression of items, while the association of the total score to age might not be large enough to warrant the calculation of separate group norms by age. The decision rule as stated in Chapter 3 that embodies the above logic follows.

> If the majority of Cramer's contingency coefficients were .20 or greater then separate percentiles by age group would be calculated. If the majority of Cramer's contingency coefficients were below .20, then a Pearson product moment correlation coefficient would be calculated for the association of chronological age ... and total PRSD raw score. If this correlation was .20 or greater then separate percentiles would be calculated for each group, on the other hand, if the correlation was less than .20 then only percentiles for the total group would be calculated.

Table 4.3 shows that only one Cramer contingency coefficient was equal to or greater than .20. Following the decision rule stated in the last paragraph, a Pearson correlation coefficient between age and total PRSD score was calculated and the results are reported in Table 4.4 below.

TABLE 4.4. SAMPLE SIZE, MEANS, AND STANDARD DEVIATIONS FOR AGE AND PRSD TOTAL SCORE, AND THE PEARSON PRODUCT MOMENT CORRELATION BETWEEN THESE VARIABLES.

Variable	N	$\bar{X}$	SD	Correlation Coefficient r_{xy}	95% Confidence Interval
Age	1,404	3.90	0.37	.161	$.122 \leq r_{xy} \leq .198$
PRSD Score	1,404	12.94	5.18		

Table 4.4 shows the correlation coefficient to be .161 between age and PRSD score. The 95% confidence interval constructed around the correlation coefficient by means of a Fisher r to Z to r transformation evidenced that the true value of the coefficients ranges from .122 to .198. Since the confidence interval does not include a value equal to or greater than .20 (as stated in the previously stated decision rule), a percentile table for all pupils irrespective of age was calculated and is presented in Table 4.5 below.

TABLE 4.5. RAW SCORE TO PERCENTILE CONVERSION TABLE FOR PRSD (N=1415).

PRSD Raw Score	Frequency	Percentile
27	1	99
26	2	99
25	5	99
24	17	99
23	19	98
22	19	96
21	29	95
20	37	92
19	59	69
18	92	83
17	85	77
16	97	71
15	94	64
14	118	57
13	110	48
12	97	41
11	110	34
10	65	28
9	83	22
8	56	18
7	65	13
6	41	10
5	32	7
4	25	5
3	12	4
2	12	3
1	4	2
0	29	1

The construct validity data set was also used to calculate two estimates of reliability for the PRSD. The Kuder-Richardson formula 20 was calculated and its result is given in Table 4.6 below along with a statistical description of the sample used for the calculations.

TABLE 4.6. KUDER-RICHARDSON 20 RELIABILITY FOR THE PRSD AND ASSOCIATED DESCRIPTIVE STATISTICS OF SAMPLE SIZE, MEAN, STANDARD DEVIATION, AND STANDARD ERROR.

Kuder-Richardson 20 Reliability	N	$\bar{X}$	SD	SE
.820	1,415	12.00	5.12	2.17

The Kuder-Richardson 20 estimate of consistency was .820 as shown in Table 4.6. The standard error of measurement using the Kuder-Richardson 20 reliability was 2.17.

The second estimate of reliability calculated was the odd-even split-half variety corrected by the Spearman Brown prophecy formula. Both the uncorrected and corrected estimates of the split-half reliability are shown in Table 4.7 below.

TABLE 4.7. SAMPLE SIZE, MEANS, AND STANDARD DEVIATIONS FOR THE ODD AND EVEN HALVES OF THE PRSD, AND PEARSON PRODUCT MOMENT CORRELATIONS UNCORRECTED FOR LENGTH AND CORRECTED FOR LENGTH BY MEANS OF THE SPEARMAN BROWN PROPHECY FORMULA.

Test Halves	N	$\bar{X}$	SD	Correlation Coefficients	
				Uncorrected for Length	Corrected for Length
Odd items	1,415	6.41	2.63	.755	.860
Even items	1,415	6.53	2.89		

Table 4.7 shows the split-half reliability to be .755 when left uncorrected for shortening the test by approximately half and the slight difference in the length of halves (12 items even and 13 items odd). The split-half reliability when corrected using the Spearman Brown prophecy formula was .860.

Overall, the construct validity study found that age progression as one element related to the readiness construct, did exist on the basis of individual PRSD items. Cramer contingency coefficients for the linear (or monotonically increasing) portion of the chi-square values ranged from .064 to .210 for 25 of the 27 items showing significance. The degree of association was judged to be minor but suggestive of age progression of PRSD items for two reasons. First, it was noted that the Cramer coefficients are somewhat stronger than equal sized Pearson coefficients

because Cramer coefficients can reach unity only if the contingency tables are infinite (in this study they were 2 x 4). Second, the age range of 3 months used on the independent variable of age limited the variation in one or both of the variables and thus caused the degree of association to be smaller. The degree of linear association between age and score status on each item would have been considerably different if a larger overall age span of subjects screened could have been included for study.

Another portion of the study using the Cramer coefficients of linear trend and a Pearson coefficient between age and total score ($r = .161$) in conjunction with an a priori decision rule, resulted in the construction of total group rather than age group percentile norms. The decision was reached on the basis of weak Cramer coefficients for linear trends for most items and a weak Pearson coefficient between age and total score. The lack of sufficient association between age and either score status on each item or total score seemed to indicate that only total group norms would be useful for the age span included in this study.

The last portion of the construct validity study concerned the estimation of reliability in terms of internal consistency. Estimates of .820, .755, and .860 were calculated for Kuder-Richardson 20, odd-even split half uncorrected, and odd-even split half corrected reliabilities

respectively. In terms of internal consistency of PRSD items, the test has good reliability.

PREDICTIVE VALIDITY

Construct validity covered in the past section and predictive validity the present topic have one, if not more, characteristics in common. Both these types of validity are classified as derived validity, as opposed to primary (or direct) validity. Findings related to primary validity require actual examination of the test to make sure its operational definition of the trait being measured is faithfully and accurately being followed. The derived (or secondary) type of validity as explored in this research is satisfied to the extent that the scores of the PRSD correlate with criterion scores that possess direct, primary validity.

Generally predictive validity is concerned with the relation of test scores to measures on a criterion based on performance at some later time. The criterion of interest in this research was school achievement. Pearson product moment correlation coefficients were calculated to determine how well one can predict school achievement (as measured by pre-test PSORT, post-test PSORT and MAT for hypotheses 2, 3, and 4, respectively) from PRSD scores. These three hypotheses are formally restated below with a presentation of the results following each.

Hypothesis 2.

Null Hypothesis: The correlation coefficient between the PRSD and the pre-test PSORT will not be greater than zero.

Symbolically: $H_o : \rho_{12} = 0$

Legend: ρ = population correlation coefficient
1 = PRSD
2 = pre-test PSORT

Alternate Hypothesis: The correlation coefficient between the PRSD and the pre-test PSORT will be greater than zero.

Symbolically: $H_2 : \rho_{12} > 0$

Legend: ρ = population correlation coefficient
1 = PRSD
2 = pre-test PSORT

This correlation coefficient with its associated 95% confidence interval for judging significance and other descriptive statistics of both variables involved in hypothesis 2 are given in Table 4.8 below.

TABLE 4.8. SAMPLE SIZE, MEANS, AND STANDARD DEVIATIONS OF THE RESULTS OF PRSD AND PRE-TEST PSORT ADMINISTRATIONS, AND THEIR PEARSON PRODUCT MOMENT CORRELATION WITH 95% CONFIDENCE INTERVAL.

Variables	Prek Pupil Classification	N	$\bar{X}$	SD	r_{xy}	95% Confidence Interval
Prekindergarten Readiness Screening Device (PRSD)--Fall, 1978	Title I	396	13.08	5.02	.487*	.409 $\leq r_{xy} \leq$ .558
Prekindergarten Saginaw Objective Referenced Test (PSORT)--Fall, 1978	Title I	396	11.10	4.96		

*p <.05.

A study of Table 4.8 above reveals that the correlation between PRSD and the pre-test PSORT was significantly greater than zero. The Pearson product-moment correlation coefficient for the 396 paired observations was .487 with an associated 95% confidence interval of .409 through .550 using Fisher's r to Z to r transformation. Thus on the basis of this sample of 396 paired observations the true correlation lies somewhere between .409 through .558 with 95% confidence. If this is true, then by squaring the limits of the interval and multiplying by 100 an estimate of the percent of variance the PRSD scores explain on the basis of the pre-test PSORT criterion scores can be obtained. Thus somewhere between 16.7 to 31.1% of the variance was shared in common

with the PRSD and pre-test PSORT scores. This level of shared variance is only suggestive that predictive validity exists, it was hoped that a moderate or larger amount of shared variance (50% or more) would be observed.

Hypothesis 3.

> Null Hypothesis: The correlation coefficient between the PRSD and the post-test PSORT will not be greater than zero.
>
> Symbolically: $H_0 : \rho_{13} = 0$
>
> Legend: ρ = population correlation coefficient
> 1 = PRSD
> 3 = post-test PSORT
>
> Alternate Hypothesis: The correlation coefficient between the PRSD and the post-test PSORT will be greater than zero.
>
> Symbolically: $H_3 : \rho_{13} \neq 0$
>
> Legend: ρ = population correlation coefficient
> 1 = PRSD
> 3 = post-test PSORT

Descriptive statistics of both variables involved in hypothesis 3 and the calculated Pearson product moment correlation along with its associated 95% confidence interval are presented in Table 4.9 below.

TABLE 4.9. SAMPLE SIZE, MEANS, AND STANDARD DEVIATIONS OF THE RESULTS OF PRSD AND POST-TEST PSORT ADMINISTRATIONS, AND THEIR PEARSON PRODUCT MOMENT CORRELATION WITH 95% CONFIDENCE INTERVAL.

Variables	Prek Pupil Classification	N	$\bar{X}$	SD	r_{xy}	95% Confidence Interval
Prekindergarten Readiness Screening Device (PRSD)--Fall, 1978	Title I	349	13.25	5.10	.383*	$.306 \leq r_{xy} \leq .471$
Prekindergarten Saginaw Objective Referenced Test (PSORT)--Spring, 1979	Title I	349	24.65	4.94		

*$p < .05$.

The correlation between the PRSD and the post-test PSORT was significantly greater than zero (Table 4.9). The Pearson product-moment correlation coefficient for the 349 paired observations was .383 with an associated 95% confidence interval of .306 through .471 using Fisher's r to Z to r transformation. Thus the true correlation lies somewhere between .306 through .471 with 95% confidence. Again estimating the shared variance in terms of a percentage, the range of explained variances falls somewhere between 9.4% and 22.2% for the association between PRSD and post-test PSORT scores. The reduced amount of shared variance between pre- and post-test PSORT relative to PRSD is probably due to the impact of the prekindergarten program in producing a negatively

skewed distribution of scores. Since the aim of the program was to foster the attainment of minimum skills stated in behavioral objective form, any students with skill levels above the stated minimums would reach the ceiling of the test and be unable to display the full breadth of their abilities. In other words, the success of the program as measured by the PSORT placed limits on the range of scores by its very minimum nature and thus it would be reasonable to expect a somewhat weaker correlation.

Before proceeding to hypothesis 4, the data from hypothesis 2 and 3 can be used to calculate a partial correlation between the PRSD and the post-test PSORT while holding the pre-test PSORT constant. The partial correlation is most valuable when the influences of one variable (initial differences on the pre-test PSORT) are to be ruled out upon the criterion behavior (the post-test PSORT) to clarify the role of the remaining variable (the PRSD). Thus the motivation for calculating the partial correlation was to determine the differential prediction of the PRSD on post-test PSORT gain scores while holding the pre-test PSORT score level constant. Table 4.10 below presents the three Pearson product moment correlations needed for calculations and the resulting partial correlation.

TABLE 4.10. SAMPLE SIZE, MEANS, AND STANDARD DEVIATIONS OF THE RESULTS OF PRSD AND PSORT ADMINISTRATIONS, AND THEIR PEARSON PRODUCT MOMENT CORRELATION AND PARTIAL CORRELATION COEFFICIENTS.

Variables	N	$\bar{X}$	SD	Pearson Product Moment Correlations		Partial Correlation[2]
				Variables		
				PSORT--Spring, 1979[1]	PSORT--Fall, 1978	PSORT--Spring, 1979
PSORT--Fall, 1978	349	11.17	4.85	.423*	--	--
PRSD--Fall, 1978	349	13.25	5.10	.383*	.520*	.211*

*$p < .05$

[1]Descriptive statistics of PSORT--Spring 1979 administration already presented in Table 4.9.

[2]The partial correlation controls for PSORT--Fall, 1978.

An examination of Table 4.10 above shows that the partial correlation between the PRSD and the post-test PSORT with the pre-test PSORT held constant was significantly greater than zero at alpha equal .05. The partial correlation of .211 explained approximately 4.4% of the common variance between the PRSD and the post-test PSORT when the effect of the pre-test PSORT was held constant.

Hypothesis 4.

Null Hypothesis: The correlation coefficient between the PRSD and the MAT will not be greater than zero.

Symbolically: $H_o : \rho_{14} = 0$

Legend: ρ = population correlation coefficient
1 = PRSD
4 = MAT

Alternate Hypothesis: The correlation coefficient between the PRSD and the MAT will be greater than zero.

Symbolically: $H_4 : \rho_{14} > 0$

Legend: ρ = population correlation coefficient
1 = PRSD
4 = MAT

The Pearson correlation coefficients central to hypothesis 4 with its associated 95% confidence intervals for the total plus two component groups are presented below in Table 4.11.

TABLE 4.11. SAMPLE SIZES, MEANS, AND STANDARD DEVIATIONS OF THE RESULTS OF PRSD AND MAT ADMINISTRATIONS, AND THEIR PEARSON PRODUCT MOMENT CORRELATIONS WITH 95% CONFIDENCE INTERVALS FOR PUPIL CLASSIFICATIONS OF ALL PREKINDERGARTENERS AND ITS TWO COMPONENTS OF TITLE I AND NON-TITLE I OR TITLE I DROPOUTS.

Variables	Prek Pupil Classification	N	$\bar{X}$	SD	r_{xy}	95% Confidence Interval
Prekindergarten Readiness Screening Device (PRSD)--Fall, 1978	All Prekindergarteners	346	13.20	5.00	.484*	.410 ≤ r_{xy} ≤ .559
Metropolitan Achievement Test (MAT)--Spring, 1980	All Prekindergarteners	346	274.42	44.24		
Prekindergarten Readiness Screening Device (PRSD)--Fall, 1978	Title I	270	13.10	4.92	.462*	.365 ≤ r_{xy} ≤ .549
Metropolitan Achievement Test (MAT)--Spring, 1980	Title I	270	273.82	45.34		
Prekindergarten Readiness Screening Device (PRSD)--Fall, 1978	Non-Title I or Title I Dropout	76	13.55	5.30	.574*	.412 ≤ r_{xy} ≤ .701
Metropolitan Achievement Test (MAT)--Spring, 1980	Non-Title I or Title I Dropout	76	276.55	40.64		

*$p < .05$.

The correlations between the PRSD and the MAT were significantly greater than zero (Table 4.11). The Pearson product-moment correlation coefficients for all 346 pupils, 270 Title I participants, and 76 non-participants were .484, .462, and .574 respectively. The three correlations had associated 95% confidence intervals resulting from Fisher's r to Z to r transformations of .410 through .559, .365 through .549, and .412 through .701 respectively. The range of explained variance on MAT by use of PRSD scores ranged from between 16.8% through 31.2%, 13.3% through 30.1%, and 17.0% through 49.1% for all pupils, Title I participants, and non-participants respectively. The overlapping nature of the correlations and the percent of variance accounted for leads to the conclusion that there was little difference on the basis of the observed sample between participants and non-participants. The obtained correlations between PRSD scores and MAT scores (received approximately 21 months later) support the supposition that some predictive validity exists. In the best of all possible worlds, it would have been hoped that observed correlations would have been greater than .7 with at least 50% shared variance.

Overall, the predictive validity study found evidence that a weak to moderate level of association existed between PRSD scores and measures of later school achievement (pre-test PSORT, post-test PSORT, and MAT). The obtained Pearson product moment correlations and their associated 95%

confidence intervals between the PRSD scores and these criterions follow: pre-test PSORT, r = .487 with interval of .409 through .558; post-test PSORT, r = .383 with interval of .306 through .471; and MAT, r = .484 with interval of .410 through .559. Thus all correlations were significantly greater than zero ($p < .05$).

Even considering that the full range of the confidence interval, these correlations fell short of a hoped for coefficient of .7 or greater, or the equivalent of 50% or more shared variance. This may be too high of a standard when considering the multitude of problems (such as distractability, shyness, negativism, and other factors interfering with rapport and test administration) associated with accurate measurement of prekindergarten level children (Anastasi, 1970, p. 473).

Beyond this consideration there were at least three other possible explanations that could account for the low correlations. First, the short length of the PRSD (27 items) played a role in reducing the size of the coefficients obtained. The focus of this test development and evaluation effort has been to produce a short test. It is generally recognized that an abbreviated form of a test is expected to have a lower correlation coefficient than the longer form of the same test.

Another factor was the Title I prekindergarten program. This program may have reduced the variability of the subjects

beyond what it might have been without such a program. It is generally recognized that all correlation coefficients are affected by a restriction in range. A reduction of the range of scores by using a homogeneous sample can result in a misleadingly low correlation coefficient. The results of the correlation of PRSD scores with MAT scores on the non-participants and dropout group is suggestive that higher correlations are possible. However, the mixed type of grouping (non-participants and partial participants) and the small number of children (N = 76) makes the results very speculative. Even within the participant group the heterogeneity of the possible participants may be further lacking due to the program leadership. The program supervisor has also been criticized a number of times for not trying harder to include even lower scoring students (on the basis of PRSD scores) into the Title I program by failing to conduct a more thorough recruitment program for participants.

The other explanation deals with a restriction of the range of differences in scores on the criterion. The post-test PSORT compared with the pre-test PSORT reflects this restriction in range. The PSORT was already described as a very program specific test in scope and thus limited by the curriculum of the program. The PSORT was further limited by its objective referenced nature to objectives usually attainable by the large majority of the pupils after

participating nine months in the program. Most pupils have been able to perform very well at the end of the program. In other words, the success of the program as measured by the PSORT brought about a negatively skewed distribution. Thus it would be reasonable to expect the PRSD to show a weaker correlation with the post-test PSORT than the pre-test PSORT.

SUMMARY

Table 4.12 below gives a synopsis of the foregoing results by test evaluation issue.

Tabl

Tes

What
rate
the

Hypo
H_0 :
H_1 :

Hypo
H_0 :
H_1 :

Hypo
H_1 :
H_1 :

Hypo
H_0 :
H_1 :

TABLE 4.12. SYNOPSIS OF TEST EVALUATION ISSUES IN QUESTION OR HYPOTHESIS FORM ALONG WITH THEIR ASSOCIATED FINDINGS.

Test Evaluation Issues[1]	Findings
What is the estimated rater reliability of the PRSD?	Intra-class correlation for an individual ($r = .789$) and the average across 19 raters ($r = .973$) is within the range of reliability felt to produce consistent results according to Guilford (1973, p. 72).
Hypothesis 1 $H_o : P_1 \geq P_2 \geq P_3 \geq P_4$ $H_1 : P_1 < P_2 < P_3 < P_4$	Fail to accept the null hypothesis and conclude that a monotonically increasing relationship is shown in the proportion of prekindergarten subjects on 25 of 27 items of the PRSD at each progressively higher chronological age interval ($p < .05$).
Hypothesis 2 $H_o : \rho_{12} = 0$ $H_1 : \rho_{12} > 0$	Fail to accept the null hypothesis and conclude that the correlation coefficient between the PRSD and the pre-test PSORT is greater than zero ($p < .05$).
Hypothesis 3 $H_1 : \rho_{13} = 0$ $H_1 : \rho_{13} > 0$	Fail to accept the null hypothesis and conclude that the correlation coefficient between the PRSD and the post-test PSORT is greater than zero ($p < .05$).
Hypothesis 4 $H_o : \rho_{14} = 0$ $H_1 : \rho_{14} > 0$	Fail to accept the null hypothesis and conclude that the correlation coefficient between the PRSD and the MAT is greater than zero ($p < .05$).

[1]Legend for hypotheses:

P_k = proportion of pupils passing an item at a given chronological age interval k (where k = 1 represents the youngest age)
ρ = population correlation coefficient
1 = PRSD
2 = pre-test PSORT
3 = post-test PSORT
4 = MAT

A review of Table 4.12 indicates that consistent results can be obtained on the basis of the observed reliability and that almost all the null hypotheses were not accepted with the exception of a couple of instances for hypothesis one. The test evaluation question involving rater reliability showed intra-class correlations within or greater than the range Guilford (1973, p. 72) termed as indicative of consistent results. Specifically, if one rater was to screen each pupil (but not the same rater for each child) then the intra-class correlation of .789 provides the best estimate of reliability using the PRSD with similarly trained raters. On the other hand, if two or more raters were to screen each child, then the intra-class correlation of .973 provided the best estimate of reliability using the PRSD with similarly trained raters.

The results related to hypothesis one testify that the null hypothesis of no age progression between scoring status on each PRSD item and age group had to be rejected for 25 of the 27 items ($p < .05$). Overall, the construct validity study found that age progression, as one element related to the readiness construct, did exist on the basis of individual PRSD items.

The remaining three hypotheses dealt with the predictive validity study. The outcomes of hypotheses two, three, and four witnessed that a correlation in excess of zero did exist. The 95% confidence intervals around the obtained

correlations showed that the coefficients of PRSD's association to the criterions ranged from .409 through .558, .306 through .471, and .410 through .559 for pre-test PSORT, post-test PSORT, and MAT respectively. This indicates that with like subjects that PRSD's shows weak to moderate predictive associations to the criterions.

CHAPTER 5: SUMMARY AND CONCLUSIONS

In this chapter, the entire study of the further development and evaluation of the Prekindergarten Readiness Screening Device (PRSD) is summarized, including the results. Additional sections of this chapter contain discussion concerning the results, conclusions, and suggestions for future research. The three measurement instruments used in the study are referred to by their acronyms in the following text.[1]

---

[1]These instruments and their associated acronyms follow: the Metropolitan Achievement Test (MAT), the Prekindergarten Readiness Sceening Device (PRSD), and the Prekindergarten Saginaw Objective Referenced Test (PSORT).

SUMMARY

T

evalua

screen

be use

and/or

typica

review

ing in

more c

tion s

validi

T

mine r

of the

one as

approp

of pre

ments.

T

seemin

childh

checkl

be due

day pr

by mea

SUMMARY

The purpose of this study was to further develop and evaluate a teacher administered measurement instrument for screening pupils for prekindergarten. The instrument would be used to identify children who, because of developmental and/or experiential problems, may be less able to meet the typical school expectations. Chapter 1 presented a review of inadequacies of available readiness and screening instruments. Inadequacies included the lack of one or more of the following: norms, description of standardization sample, data on test reliability, and/or data on test validity.

The specific purposes of this study were (1) to determine rater reliability, (2) to explore the age progression of the percent passing each item of various age groups as one aspect of construct validity (and in the process develop appropriate percentile norms), and (3) to obtain estimates of predictive validity with a number of achievement instruments.

The literature reviewed in Chapter 2 covered two seemingly unrelated topics - historical changes in early childhood education's mission and the error associated with checklists. The linkage between the topics was shown to be due to the compensatory educational nature of present-day preschool programs that require screening (sometimes by means of checklists) of potential participants. Early

childhood education as an organized enterprise in the United States has a short history. The first such endeavors started in the 1920's as nursery schools attached to universities and colleges for research purposes. Their curriculum consisted of habit training and the promotion of physical health. Federal funding during the depression years and then again during World War II increased the number of nursery schools but changed little their curriculum other than an increase of an emphasis on socio-emotional growth in the 1940's. During the 1960's, a combination of academic and socio-political forces finally lead to massive federal funding of compensatory education preschool programs designed to help prevent school failure among lower-lower-class children. The curriculum of these preschools emphasized cognitive growth and other abilities necessary for success in schools while drawing curriculum elements from the former nursery school movement. Present-day preschools enjoy a great variety in the composition of curriculum, but for the most part focus upon the following four areas of development: socio-emotional, perceptual-motor, cognitive, and language. Any readiness screening device for preschool education should be broad enough to cover these areas of child development.

Screening for preschool may be accomplished by means of a checklist as one way to measure readiness skills. While checklists at first glance appeared to be a simple

measurement tool, errors of measurement were possible from a number of sources. The personality of the rater, the scale itself, the nature of the action (or trait), and the opportunity afforded the rater were all classified as common sources of error. Specific suggestions from the literature were offered in Chapter 2 to combat these errors under the broad headings of clearer definition of action rated and provisions for training raters.

The School District of the City of Saginaw's Title I Prekindergarten program served as the setting for the study. Both prekindergarten staff members and potential pupils served as subjects. Data from one or more of the 1978-79, 1979-80, and 1980-81 school years served as the basis for the three major developmental thrusts of the study - 1) determine rater reliability, 2) estimate construct validity, and 3) assess predictive validity. The PRSD, the PSORT, and MAT were described in Chapter 2 with reliability estimates of .761 and .909 offered for PSORT and MAT respectively.

The rater reliability study consisted of raters scoring two videotaped testing sessions. A short training session was held prior to scoring two children on videotape. Intra-class correlations were calculated. Also percent agreement of these raters were compared to the original rater and items identified as difficult to score during the scoring sessions.

The total set of the PRSD results from the three consecutive school years mentioned above was employed in the construct validity study. This data set was used to study the construct validity of the items by means of a test for trend in ordered contingency tables. Chronological age intervals crossed with pass and fail on PRSD formed the cells for the ordered contingency table for each item. The level of significance was alpha = .05. Percentile norms either for the entire group or each age interval group were to be developed on the basis of decision rules relating to the results of the construct validity study.

In the predictive validity study the results of September and October, 1978 administration of the PRSD served as the test scores (or the X variable) to predict the criterion. The results of the PSORT administration of October, 1978 (pre-test) and April, 1979 (post-test) for the Title I Prekindergarten program plus the MAT administration of April, 1980 served as the achievement criterion measures. The Pearson product-moment correlation coefficients were calculated for each of the criterions. To test whether these correlations were significantly different than a zero correlation, a Fisher r to Z transformation was used. The level of significance was alpha = .05.

Reliability Study Results

The reliability study sought to estimate the rater reliability for the entire test and explore the rater agreement to the original rater by item. The intra-class correlation for an individual rater was .789 and the average for the group of raters in this study was .973. The percent of agreement to the original rater ranged across items from 5.3 to 100% and 21.1 to 100% for pupils A and B respectively. Only item 14 (walks backward) obtained consistently low percent agreements of 10.5 and 21.1% for pupils A and B respectively. When a criterion of less than 75% percent agreement was used to single out the other remaining items where agreement was low, the following items were evident: 1 (say age) 68.4%, 17 (draw cross given model) 42.1%, 24 (draw diamond given model) 52.6%, 25 (tell nursery rhyme, song, or poem) 10.5%, 26 (say yellow jello) 63.2%, and 27 (more than one word response) 5.3%. After the scoring session a discussion was held concerning particularly hard to score items and/or suggestions about improved procedures related to testing, only low scoring items 24 (draw diamond given model) 52.6% and 25 (tell nursery rhyme, song, or poem) 10.5% received comments. These comments were "It was hard to demonstrate tracing around template"(item 24) and "You should reword and simplify item 25 by asking, 'Can you tell me a song or poem?'"

Construct Validity Study Results

The construct validity study explored the age progression of the percent passing each PRSD item as one aspect of the readiness construct. This information along with Cramer's contingency coefficients for each item's linear component and a Pearson product moment correlation between total score and age was used to determine if age interval or entire group percentile norms should be developed. In addition, the same data set was used to estimate how consistent the PRSD items were in providing reliable measurement by means of the Kuder-Richardson 20 formula and the split-half reliability corrected by means of the Spearman Brown prophecy formula.

The null hypothesis of no age progression in the percent passing each item was rejected for 25 of the 27 items ($p < .05$). Only results of item 19 (throw a ball five feet) and 27 (more than one word responses) failed to reject the null hypothesis. The Cramer coefficient ranged from .064 to .210 (average Cramer coefficient = .136) for the 25 items showing significance. The majority of the Cramer coefficients were clustered between .13 and .18.

Since a majority of the Cramer coefficients were below .20, a Pearson product moment correlation coefficient was calculated between the PRSD and age with the resulting coefficient equal to .161. On the basis of these somewhat weak correlation coefficients, percentile norms for PRSD

scores were calculated for all students irrespective of age groupings.

The estimate of reliability in terms of internal consistency of the test were calculated to be .820 and .860 for Kuder-Richardson 20 and odd-even split half reliability corrected by Spearman Brown prophecy formula respectively.

Predictive Validity Study Results

The predictive validity study estimated the strength of relationship between the PRSD and future measures of school achievement (PSORT and MAT). The correlations between PRSD and the criterion instruments were estimated by means of 95% confidence intervals. These intervals were .409 through .558, .306 through .471, and .410 through .559 for pre-test PSORT, post-test PSORT, and MAT respectively. The correlation related to the MAT was recalculated on the basis of subjects classified either in the Title I prekindergarten program or in a combined category of non-Title I participants or Title I dropouts. The 95% confidence intervals around each of the resulting coefficients were .365 through .549 and .412 through .701 respectively.

This concludes a short review of the results related to the PRSD's predictive and construct validities and reliabilities encompassed by this research. A discussion of these results follows.

DISCUSSION

The results of this further development and evaluation of the PRSD should be qualified. The three connected studies used data gathered over a period of three years from a realistic setting with some, but by no means total control of events being possible. Events, such as a teacher strike which delayed screening during the 1979-80 school year by a few weeks, staff changes within the prekindergarten program, unwillingness of a prekindergarten staff member to follow directions, and the refusal of the prekindergarten program supervisor to use additional means to publicize the program to increase the number of children screened, had some effect on the internal validity of the study. These events may have altered the results from what they could have been if controlled. Thus, while the results are generalizeable, the true statement of power of the PRSD still must wait for further controlled studies.

A discussion of the results related to the three psychometric characteristics--reliability, construct validity, and predictive validity follow. All three characteristics have something to add to the description of overall measurement quality of the PRSD.

Reliability

Reliability estimates for psychological and educational tests can be thought of as statistical indices of measurement quality. Reliability is a necessary but not sufficient condition to assess the overall quality of a test. Even with a highly reliable measurement instrument, an author who assesses trivial or irrelevant aspects of a phenomenon has done nothing more than precisely measure some unimportant aspects of the phenomenon. Nevertheless, only to the extent that test scores are reliable can they be useful for their intended purpose.

In the case of the PRSD, the primary issue of reliability centers upon how accurately a rater or multiple raters can assess the performance of pupils. The PRSD raters on an individual basis do a satisfactory job of bringing about consistent and objective results ($r = .789$) or when multiple PRSD raters will screen a single pupil ($r = .973$).

Additional steps are needed to improve rater reliability and/or analysis of rater behavior. More "dry" runs with real subjects or viewings of videotapes and greater discussion of differences in ratings afterwards seem to be means to further help the present group of raters or any group that shows some substantial disagreement in scoring. In addition, the standard of comparing the original rater with the final group of raters may have been inappropriate

because of some technical qualities of the videotape (sound quality, clarity of picture, and rater's line of sight). A better standard might have been the percent incorrect of videotape raters across subjects with no reference to the original rater. Under this system, items that consistently show a mid-range of percent incorrect (35% to 65%) across a clear majority of subjects seem the best candidates for further revisions in the directions (or demonstrations) given to the pupils and/or the scoring instructions to the raters. None of the results by item showed a mid-range percent incorrect for both subjects tested. Any further definitive answer has to wait until more than two pupils can be involved in any future rater reliability study. However, possible candidate items for changes in directions (or demonstrational procedures) provided to the pupil include the following items along with the suggested change: item 14-clearer demonstration of walking backwards with more verbal description of what is involved; item 24-consistent demonstration of how to trace around diamond template by picking up pencil after each straight line; item 25-simplify by asking, "Can you tell me a song or poem?" Possible candidate items for changes in scoring instructions include the following items along with the suggested changes: item 18-further specify forward hop is correct and an in place hop is incorrect; and item 19-allow three throws and if two of three throws meet the criteria then score as correct. These changes could

reduce leniency errors of raters that may result when raters feel the test instructions for scoring procedures are less than adequate.

In addition, coefficients of test reliability for the PRSD were calculated on the basis of the data set used for the predictive validity study. Coefficients of internal test reliability indicate how similarly the pupils perform on different, but supposedly equivalent, items. The coefficient of rater reliability, as discussed earlier, focused on rater agreement; while coefficient of test reliability were aimed at examinee performance. Both the Kuder-Richardson 20 formula ($r = .82$) and the odd-even split half coefficient corrected by the Spearman-Brown formula ($r = .86$), provided evidence that the results of pupils are satisfactorily consistent across items. These correlations show that approximately 67 to 74% shared variance was accounted for by each of the techniques to estimate internal consistency between one portion of the PRSD and another portion of the same instrument.

Construct Validity

The PRSD, in the educational context of this study, seems to possess a satisfactory estimate of reliability (both in the rater and internal test reliability sense) to bring about consistent results. However, what about its validity (or the other half of the index of measurement

quality)? Validity, or the ability of a test to do the job for which it is intended, takes on many connotations. The importance of each type of validity varies with the purpose of the test. The PRSD is intended to predict youngsters who will have trouble with future school achievement and thus is oriented to measure the precursors of academic skills. While predictive validity is of primary importance, results were gathered to substantiate two other types of validity. Evidence relating to face/content validity of PRSD items in terms of research done by other investigators has already been presented in Appendix G. The findings related to construct validity need to be analyzed before turning to a discussion of the predictive validity study results.

The construct validity study found that age progression, as one element related to the readiness concept, did exist for 25 of the 27 PRSD items. The degree of linear association due to age related to ability to pass a PRSD item could be judged to be weak to moderately weak for the 25 items. These results seem similar to the results of other investigators presented in Appendix G.

There are at least two reasons why the degree of linear association between age and scoring status on each PRSD item was not stronger. First, the amount of variation observable on the PRSD was limited by using subjects essentially spanning a one year age period. Without enough variation, it was to be expected that a lowered degree of

association would be found in this study than another study employing subjects with a greater range of ages. The experience gained from this research suggests that age progression can better be investigated when subjects span three or four years in age. A definitive test of age progression does seem to require a broader age range of subjects than one year.

Secondly, more age intervals would also seem more appropriate. The present research was somewhat limited to the three month age interval due to an attempt to keep a similar number of pupils in each age interval. However, the observed differences would no doubt have been greater if more than four age intervals could have been employed practically.

The correlation between the total PRSD score and age was undertaken to explore if the age progression shown by youngsters on one item would tend to be additive. Since the correlation of total score and age was approximately the same size as a single item, this result may suggest that readiness skills as measure by the PRSD cover a broad area of skills that do not appear at this age to build upon each other. A factor analysis might be helpful to shed further light on this issue of dimensionality of skills measured and further define the construct of readiness embodied in the PRSD. However, the high estimates of internal consistency tend to suggest a single factor explanation to PRSD's construct of readiness.

Predictive Validity

The PRSD predicted the achievement criterions of the pre-test PSORT, the post-test PSORT, and MAT to a weak to moderate degree. The objective referenced test construction of the PSORT may have been responsible in part for the lower than expected correlation of the PRSD with the post-test PSORT. Objective referenced tests generally are tailored to measure objectives specific to the educational program. If a program is successful, the participants' post-test scores on objective referenced tests are usually very close to perfect scores. In such a situation, most children would have near perfect scores and be unable to show any greater ability because they were topping-out on the measurement instrument, like the PSORT. Thus, the objective referenced nature of the PSORT coupled with a fairly successful program may be a portion of the reason why the correlation of PRSD and post-test PSORT is lower than the correlation of PRSD and pre-test PSORT.

A series of conclusions follow.

CONCLUSION

The specific purpose of this study was to further develop a teacher administered measure for screening prekindergarten pupils. The three major thrusts of this developmental research effort were: 1) to determine rater reliability, 2) to estimate one aspect of construct validity (age progression) and develop percentile norms, and 3) to assess predictive validity. The following conclusions are intended to accurately summarize the major findings of this study.

- Rater reliability of the PRSD was shown to be substantial enough to insure that consistent measurement was taking place either by a single rater or a group of raters rating the same pupil. Particular recommended changes in the directions to pupils and the scoring procedures were offered to help improve this reliability even more.
- All except two of the PRSD items show an increasing percentage of pupils passing the items as chronological age increases.
- A percentile norm table was developed to help assist teachers in explaining the results of the PRSD to parents in terms of their child's relative rank to other Title I children in the norming population.
- The PRSD appears to be a weak to moderately weak predictor of achievement on the PSORT and MAT.
- Estimates of internal consistency of the items appear to show adequate Kuder-Richardson 20 and split-half reliabilities.

IMPLICATIONS FOR FURTHER RESEARCH

The following implications for further research are offered.

- Predictive validity studies with other criteria as grades, parental report, longer developmental instruments, and other achievement measures should be undertaken. It seems prudent to limit the time span to the minimum of six months or less when conducting such studies in the future.

- Factor analysis of the PRSD is warranted to determine if unidimensional factors exist that would make subtest analysis feasible. The determination of subtest scores could provide additional diagnostic information related to areas of developmental lag.

- The nomological network of the construct of readiness as typified by the PRSD should be explored in relationship to other factors. Some factors to include would be the following: attention span of child, experience data reported by parents, developmental history data, social economic status of family, and birth order.

- Since rater reliability was good to excellent on most tasks, further scaling of the correctness incorrectness continuum of the PRSD from a 2-point to a 3-point or 4-point scale to yield even more differentiating data in the same amount of time should be explored.

- Age progression studies on the PRSD should be carried out that span 3 to 4 years to further demonstrate the utility of the items and further determine what would constitute the proper span of age for the norming tables.

APPENDICES

APPENDIX A

TABLE A.1.
COMPARISON OF ACTUAL ADMINISTRATION TIME AND GRADE LEVEL(S)/AGE(S) FOR SIXTEEN READINESS INSTRUMENTS PRESENTED IN THE EIGHTH MENTAL MEASUREMENT YEARBOOK

Test Name	Grade Level(s)/ Age(s)	Actual Administration Time in Minutes
The APELL Test: Assessment	ages 4-5-7	40
Academic Readiness Scale	first grade entrants	5-10
Analysis of Readiness Skills Reading and Mathematics	kindergarten - 1	30-40
Cognitive Skills Assessment Battery	prekindergarten	20-25
Hess School Readiness Scale	ages 3.5-7.0	8
Initial Survey Test	first grade entrants	120-150
Jansky Screening Index	kindergarten	15-20
The MacMillan Reading Readiness Test, Revised Edition	first grade entrants	75-90
Metropolitan Readiness Tests, 1976 Edition	first half kdgn. (Level 1)	105
	second half kdgn. & first grade entrants (Level 2)	110
Murphy-Durrell Reading Readiness Analysis	first grade entrants	60
PMA Readiness Level	kindergarten - 1	65
PreReading Expectancy Screening Scales	first grade entrants	25-35
Riley Preschool Developmental Screening Inventory	ages 3-5	3-10
School Readiness Survey, Second Edition	ages 4-6	15-30

Table A.1. (Cont'd.)

Test Name	Grade Level(s)/ Age(s)	Actual Administration Time in Minutes
School Readiness Test	kindergarten - 1	60
Thackray Reading Readiness Profiles	ages 4 years, 8 months to 5 years, 8 months	70

APPENDIX B

May 13, 1976

Dr. Ralph Scott
Iowa Test of Preschool Development
Go-Mo Products
1171 7th Street
Des Moines, IOWA 50314

Dear Dr. Scott:

For the past several years, we have been searching, unsuccessfully, for a norm-referenced test to include in our pre-school evaluation component. All our instruments are locally developed and objective referenced. Our search for a norm-referenced test came to a halt about two months ago when we received a brochure outlining your Iowa Test Of Preschool Development. It appeared to be just what we were looking for. We were even more encouraged when a specimen set arrived.

Our initial encouragement faded quickly, however, when we read chapter two in the test manual. We discovered that blacks and other non-whites were excluded from your norming sample even though some 200 were originally selected. We had a difficult time understanding how the exclusion of these children was going to further the goal of an "integrated society," which you mentioned on page 8. By excluding all blacks from your norming sample, you seem to be implying that they will lower the test's norms. It seems to us that your procedure will produce an invalid standard, the same as if you would have excluded all poor rural children, or some other group. We don't believe that there should be separate norms for separate groups of children, but we do believe that any norm should include a representative cross-section of the entire population. This is not the case with the I.T.P.D. Since our Preschool Program, which is funded through Title I, is comprised almost entirely of black, inner-city children (over 400 of them), we feel the norms of the I.T.P.D. would be inappropriate for us.

We have discussed this norming problem at length and cannot understand, try as we might, why you used this norming process. To be completely honest, we would feel uncomfortable using a test that excluded all blacks from the norming sample. And we think most evaluators and program directors of inner city preschools would have a similar reaction.

Now that we have spelled out our reactions to your test, we would welcome a response from you. We are sorry that your test cannot be used in the Saginaw Public Schools, but it is just not applicable to our situation. The specimen set of the test is being returned with this letter.

Sincerely,

Barry E. Quimper, Supervisor
Program Evaluation

Michael Manley
Reporting Specialist/Program Evaluator

kjk

UNI

UNIVERSITY OF NORTHERN IOWA · Cedar Falls, Iowa 50613

EDUCATIONAL CLINIC
Department of Educational
Psychology and Foundations
AREA 319 273-2648

May 20, 1976

Mr. Barry E. Quimper, Supervisor
Program Evaluation
School District of the City of Saginaw, Michigan
Administration Building
550 Millare St
Saginaw, Michigan 48607

Dear Mr. Quimper

Mr. Stu Duncan of GoMo Industries has relayed to me your May 13 letter for reply. I sincerely regret that it is not possible to talk about the significant and practical questions you raise. The most outstanding question is one which troubled me as we proceeded to refine the ITPD: indeed, the decision to use only white subjects in the norming group was a matter over which we agonized.

For better or worse, we decided to use only white children in the norming (although hundreds of black children were tested, and in the Home Start project the majority of participating youngsters were black.) The chief reason for this was a statement by James Coleman, author of the Coleman Report, to the effect that there is an existing achievement gap between most blacks and most whites and it is the job of the school to close that gap. In other words, to deal with the very real problem of the average black graduating from high school with seventh grade arithmetic and ninth grade reading competencies, we decided to use white norms as our early assessment guides and to promote parental counseling accordingly.

We are now fortunate in having long-term results of the Home Start project which used our strategy. Briefly, our design required comparing third grade scores of Home Start children on the Iowa Test of Basic Skills with the comparable scores of their older and non-Home Start siblings. This allowed us to control for family background factors. These results will be published in the October, 1976, issue of Psychology in the Schools. I have been told that these results are among the most encouraging in the nation. It may also be of interest to you that EARLY YEARS journal has published a book which breaks down early enrichment activities into the 11 readiness areas of the ITPD, permitting easy translation of test results into practical individualized enrichment.

I hope that this will be somewhat helpful. Frankly, I share your sentiments about the problems of norming and hope you will let me have any other questions that come to mind.

Yours sincerely

Ralph Scott, PhD
Director, Educational Clinic

PS

	BLACKS			WHITES			BLACKS & WHITES		
	t-test	Mean		t-test	Mean		t-test	Mean	
		Home Start	Non-Home Start		Home Start	Non-Home Start		H.S.	N.H.S
Vocabulary	.3	2.4	2.6	.4	2.8	2.9	.5	2.5	2.7
Reading	.4	2.2	2.4	.2	3.0	2.9	.3	2.5	2.5
Spelling	.05	2.6	2.6	.4	2.9	2.7	.1	2.6	2.6
Capitalization	2.3*	2.8	1.7	.8	3.4	2.7	2.3	3.0	2.0
Punctuation	.05	2.3	2.3	.8	2.5	2.9	.3	2.4	2.5
Usage	1.2	2.2	1.9	.2	2.9	2.8	.9	2.4	2.2
Total Language	2.2*	2.3	1.9	.2	2.8	2.7	1.5	2.4	2.1
Maps	2.2*	2.6	1.8	.05	2.7	2.7	1.9	2.7	2.1
Graphs	1.0	2.7	2.1	.4	2.9	2.7	1.1	2.8	2.3
Reference	2.7**	2.8	2.0	.6	3.6	3.1	2.2	3.0	2.3
Total Word Study	2.6**	2.7	1.8	.4	2.7	2.9	2.1	2.7	2.1
Concepts	1.2	2.8	2.3	.5	3.1	3.3	1.8	2.9	2.6
Problems	3.8**	2.9	1.7	.1	3.1	3.0	3.1	2.9	2.1
Total Mathematics	2.9**	2.7	1.8	.1	3.1	3.1	2.4	2.9	2.2
Composite	2.2*	2.3	1.9	.3	2.9	3.0	1.2	2.4	2.2

PREKINDERGARTEN

READINESS

SCREENING

DEVICE

DIRECTIONS FOR ADMINISTERING PREKINDERGARTEN READINESS SCREENING DEVICE (PRSD)

- Read these thoroughly before testing -

General Directions

Make every attempt to administer all items to children being tested.

All directions that are for teacher use only will be in parentheses ( ) and are not to be read to the child. This test is to be administered on a one-to-one basis. During the administration of PRSD, an effort should be made to use the words printed in the directions.

Take as much time as you think is necessary to administer the test to each child. You should be able to handle all items in one test session, since the test is relatively short.

Since you are to record as correct or incorrect the responses, your directions for each test item will have a section describing what the correct or acceptable responses are so that you can appropriately mark your scoring sheet.

Familiarize yourself with all testing materials before administering the test. Make sure that you have a complete set of manipulative materials and flash cards before beginning the test. In addition, practice read all the specific directions for administering the test items at least once before testing begins. This practice should help you more quickly score each response since you will be more knowledgeable of acceptable responses as well as more aware of the flow of test activities.

Test Administration/Scoring

All responses are to be accepted, i.e., the content of a child's responses should be accepted by the test administrator. Ways (words and gestures) of accepting the content (whether correct or incorrect) should be limited to smiles, nods, or such positive phrases as "thank you," "okay," "fine," etc.

The person administering the test should, however, make sure that the child clearly understands the task that s/he is to perform, i.e., s/he should be told how to respond to the question and this may involve correcting the child.

The person administering the test immediately scores each item at the time of recording responses on the answer sheet. A separate machine readable answer sheet will be provided to those administering the test and it should be marked as the test is administered. Use these symbols for scoring:

A for correct responses, and

B for incorrect or no responses

Identifying Information

On the machine readable answer sheet print the child's name, last name first, then first name, etc., in the boxes provided on the scoring sheet. Be sure to blacken in the letter boxes below the child's name.

Also print and darken in the boxes corresponding to the birth date (month and year) and sex of each child in the appropriate boxes provided on the answer sheet. In column number 1 use the

following code numbers and definitions to indicate the racial/ethnic designations for each child.

Code	Definitions*
1	AMERICAN INDIAN OR ALASKAN NATIVE OR NATIVE AMERICAN - A person having origins in any of the original peoples of North America.
2	WHITE, NOT OF LATINO OR HISPANIC ORIGIN - A person having origins in any of the original peoples of Europe, North Africa, the Middle East, or the Indian subcontinent.
3	LATINO OR HISPANIC - A person of Mexican, Puerto Rican, Cuban, Central or South American, or other Spanish Culture or origin, regardless of race.
4	BLACK, NOT OF LATINO OR HISPANIC ORIGIN - A person having origins in any of the black racial groups.
5	ASIAN OR PACIFIC ISLANDER - A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Pacific Islands. This area includes, for example, China, Japan, Korea, the Philippine Islands, and Samoa.

*Race/ethnic designations as used by the Michigan Department of Education do not denote scientific definitions of anthropological origins. For the purpose of this report, a student may be included in the group to which he or she appears to belong, identifies with, or is regarded in the community as belonging. However, no person should be counted in more than one race/ethnic group.

Print the name of the school and your name (instructor) on the lines provided.

Setting/Equipment

Nearly any room (unused classroom, office reception area, etc.) relatively free of distractions can be used for test administration. Two tables and at least four chairs should be

available so while the teacher tests the child in one corner of the room the aide may interview the accompanying parent in the other corner.

In all cases the person administering the test must first establish some rapport with the child before beginning testing. Parents are welcome in the room but the person administering the test should not involve the parent in presenting test items or interpreting responses.

PREKINDERGARTEN READINESS SCREENING DEVICE (PRSD) TEST ITEMS

Remember, all statements in parentheses ( ) are intended for your use and are not to be read to the child.

Sample 0. SAY, "Let's play some question and answer games together. Okay. The first thing I'd like for you to tell me is your name. What is your name?"

(Pause for child's response.)

SAY, "(Insert child's response), that is a nice name."

(Do not score this sample item.)

REMEMBER IN RECORDING ANSWERS ON YOUR ANSWER SHEET YOU SHOULD CODE A = CORRECT RESPONSE, B = INCORRECT OR NO RESPONSE.

1. SAY, "What is your last name?"

(Pause. Repeat the child's response and mark your scoring sheet according to the acceptable response given below.)

Acceptable Response

--Any response mentioning the child's last name. You should probe if s/he only gives his/her first name.

2. SAY, "How old are you?"

(Pause. If the child holds up fingers to indicate his/her age then ...)

SAY, "Can you tell me how many that is?"

(Pause. Repeat the child's verbal response and mark the answer sheet according to the response categories given below.)

Acceptable Response

--Responds verbally with correct age.

---

3. SAY, "Now can you do something else for me. Can you put your finger on your neck. Show me where your neck is."

(Pause. Watch for one of these acceptable responses given below and mark your answer sheet accordingly.)

Acceptable Responses

--Points to neck.
--Places fingers on neck.

4. SAY, "Tell me something else."

(Teacher points to child's elbow while asking the question below.)

SAY, "What do you call this? It's a part of your arm. What is it called?"

(Pause. Repeat the child's response and mark your scoring sheet according to the acceptable response given below.)

Acceptable Response

--Elbow.

---

5. SAY, "I would like to ask you a question. What are your ears for? Can you tell me what you use your ears for?"

(Pause. Repeat the child's response and mark your scoring sheet according to the acceptable responses given below.)

Acceptable Responses

--Listen to music, people, TV, etc.
--Hear music, people, TV, etc.

6. SAY, "Okay. We're going to get the blocks out and play a few games."

(Teacher displays 4 blocks - yellow, green, black and blue on table before the child.)

SAY, "Do you see all the blocks on the table?"

(Pause for an affirmative response. Then the teacher holds the blue wooden cylinder so the child can see it.)

SAY, "Now, what I would like you to do is give me one this color."

(Pause. Mark your answer sheet according to the acceptable responses given below.)

Acceptable Response

--Blue block is handed to teacher.

---

7. SAY, "Let's play another game with the blocks."

(Teacher should have the four different colored blocks on the table before the child. The blue wooden cylinder should be returned to the box.)

SAY, "Show me the green one."

(Pause. Mark your answer sheet according to the acceptable response given below.)

Acceptable Response

--Green block is pointed to or picked up.

8. SAY, "Let's find out something else about the blocks."

(All four blocks should be on the table. Teacher points to yellow block on the table before the child.)

SAY, "What color is this?"

(Pause. Repeat the child's response and mark your answer sheet according to the acceptable response given below.)

Acceptable Response

--Yellow.

---

9. (All four blocks should be on the table. Teacher points to the black block.)

SAY, " What color is this?"

(Pause. Repeat the child's response and mark your answer sheet according to the acceptable response given below. Blocks should be removed from the table.)

Acceptable Response

--Black.

10. SAY, "Now can we try some counting. Can you count? Count to five for me."

(Pause. Mark your answer sheet according to the acceptable response given below.)

Acceptable Response

--One, two, three, four, five.

---

11. SAY, "Now I'm going to put these blocks down here again."

(Use the five blocks - blue, yellow, green, black and red.)

SAY, "Give me four blocks."

(Pause. Mark your answer sheet according to the acceptable response given below.)

Acceptable Response

--Child hands teacher four blocks.

---

12. SAY, "Now I'm going to play a trick on you. I'm going to take some of these blocks and then I want you to tell me how many are on the table. How many do you see?"

(Teacher places three blocks on table and continues to hold the other two blocks. Pause. Repeat the child's response and mark the answer sheet according to the acceptable response given below.)

Acceptable Response

--Three.

13. (Teacher sets all five blocks on the table in front of the child.)

SAY, "Now, give me <u>all</u> the blocks?"

(Pause. Mark the answer sheet according to the acceptable response given below.)

<u>Acceptable Response</u>

--All five blocks are handed to the teacher.

SAY, "Now I'm going to put them back into the box.

(Teacher places the five blocks in the box.)

---

14. SAY, "We're going to do some other things. I'm going to push my chair back and would you watch me walk because I'm going to see if you can walk just like this. Okay?"

(Teacher walks backwards toe to heel at least six steps or more, while keeping the toe and heel touching or not more than 1 inch apart.)

SAY, "Can you walk like that?"

(Pause for the child to demonstrate. Mark the answer sheet according to the acceptable response given below.)

<u>Acceptable Response</u>

--Child walks backwards toe to heel two or more steps with heel within 1 inch of toe.

15. SAY, "Now before you go back to your chair. I want you to listen carefully and do what I say. Would you clap your hands one time, walk around the chair, and sit down."

(Pause for child to demonstrate. Mark the answer sheet according to the acceptable response given below.)

Acceptable Response

--Child carries out the three part command in the specified order -- claps hands once, walks around chair, and sits down.

---

16. (After the child has completed his/her attempt, please have the child sit down.)

SAY, "Can you tell me something now? What are books for?

(Pause. Repeat same questions if no response or if child gives an incorrect response then ask.)

SAY, "Can you do anything else with books?"

(Pause. Repeat child's responses and mark the answer sheet according to the acceptable responses given below. If the child gives one or more of these responses the item is scored as correct.)

Acceptable Responses

--Read them.
--Color in them.
--Look at them.
--Tell stories with them.
--Learn things from them.

17. SAY, "I'm going to show you something else I have here."

(Teacher places the cross flash card on the table in front of child with one of the lines pointing directly at the child.)

SAY, "And I'll give you a piece of paper. Do you like making things on paper?"

(Pause.)

SAY, "I'm going to give you this crayon. I'd like you to make me something that looks like the lines on the card."

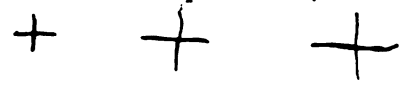
(Teacher points to the black lines and traces them with his/her finger.)

SAY, "Would you draw that?"

(Pause for child to complete the drawing. Mark the answer sheet according to the response categories given below.)

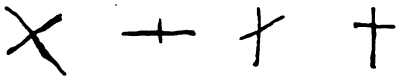
Acceptable Response

--Two lines, clearly crossing each other and roughly straight, no rotated plane (see illustrations below).

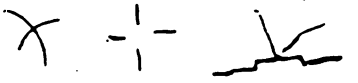


Incorrect Responses

--Two lines cross at quarter point or less, 1 line twice the length of the other, cross in rotated plane (see illustrations below).



--One or two lines, no crossing, curved lines, scribble, no resemblance to cross (see illustrations below).



(Teacher places crayon and flash card into the box.)

18. SAY, "Now we're going to get up again, it's kind of fun to get up and move around isn't it?"

(Pause.)

SAY, "Watch me and see what I'm doing."

(Teacher demonstrates a one-foot hop by hopping at least six times.)

SAY, "Now can you hop like I did? Try it."

(Pause. Mark the answer sheet according to the acceptable responses given below.)

Acceptable Response

--Child hops on one foot for two or more hops.

---

19. (Teacher takes multi-colored ball from the box.)

SAY, "Now you can stand right here. Do you see this pretty ball I have?"

(Pause and hand the ball to the child.)

SAY, "Would you hold that and I'm going to move over this way."

(Teacher points to an area five feet from the child and walks to that spot.)

SAY, "Now I'm all set so would you throw the ball to me."

(Pause for the throw. Mark the answer sheet according to the acceptable response given below.)

Acceptable Response

--Child throws ball within reach of teacher (any type of throw is acceptable).

20. (Return the ball to the box and obtain the flash cards with the triangle, circle, and line. Position the cards in front of the child pointed at him/her from left to right in the following order and positioning: line pointing at the child, circle, and triangle with top angle pointed away from child.)

SAY, "We have just a few more things to do. Now I'm going to show you some other cards I have. I'm going to put these cards like this. And I'm going to ask you some questions about them. Are you ready?"

(Pause for affirmative response.)

SAY, "Can you show me which one of these is most like a wheel?"

(Pause for the child to indicate his/her response. Mark your answer sheet according to the acceptable response given below.)

Acceptable Response

--Circle.

---

21. SAY, "Fine, look at the cards again. This time tell me which one looks most like a tent?"

(Pause for the child to indicate his/her response. Mark the answer sheet according to the acceptable response given below.)

Acceptable Response

--Triangle.

22. SAY, "Look at the cards again. This time show me which one looks most like a stick?"

(Pause for the child to indicate his/her response. Mark the answer sheet according to the acceptable response given below. Teacher replaces cards in the box.)

Acceptable Response

--Line

---

23. SAY, "Alright, now we have a question about how big things are. Which is bigger, a tree or a flower?"

(Pause. Repeat child's response and mark the answer sheet according to the acceptable response given below.)

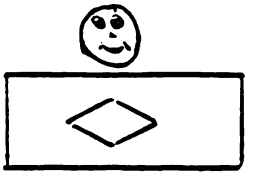
Acceptable Response

--A tree.

24. (Teacher should obtain the diamond template, crayon, and a piece of paper.)

SAY, "I have something else I'd like you to do with a shape. I'm going to give you this piece of paper, shape, and a crayon."

(The paper and template should be laid lengthwise in front of the child as illustrated below.)



SAY, "I want you to draw around the shape like this."

(Teacher traces around the shape with his/her fingers while holding the template in place.)

SAY, "Would you use the crayon to do that for me?"

(Pause for the child to complete his tracing. Mark the answer sheet according to the acceptable response given below. Teacher returns the crayon and template to the box.)

Acceptable Response*

--Outline of the diamond no more than 3/4 of an inch from the edge of the template.

---

*The standard was determined by first measuring the deviations accepted as correct from previous test administrations, finding the mean of these deviations and, finally, dividing the mean by one-half and adding it to the mean. This standard closely matches those quoted by other school readiness researchers.

25. SAY, "Now I'd like you to do something for me. This is going to be lots of fun. Do you have a favorite nursery rhyme, poem, or song?"

(Pause for the answer. After determining which rhyme, poem, or song, if any, proceed. If child does not know one proceed to next question.)

SAY, "Could you tell me that nursery rhyme or song?"

(Pause for child's response.)

SAY, "Is there any more?"

(Pause for child's response. Mark the answer sheet according to the acceptable response given below.)

Acceptable Response

--Child correctly recites at least 2 complete lines of rhyme, poem, or song.

---

26. SAY, "Oh, I liked the way you said that. I have one more thing I'd like you to say and it sounds funny. Can you say yellow jello?"

(Pause. Mark the answer sheet according to the acceptable response given below.)

Acceptable Response

--Child must pronounce the j and y clearly.

27. (Score whether the child tended to engage in conversation-like responses during the testing session. Mark the answer sheet according to the acceptable response given below.)

Acceptable Response

--Child responded to questions with more than just one word responses to a clear majority of the items requiring a verbal response. This item is scored as either correct (A) or incorrect (B).

SAY, "Well you did a very nice job. In fact, you have done such a good job for me I have a little surprise for you.

(Teacher hands the child a surprise, e.g., candy, ring, picture, etc. Be sure all items have been recorded and identifying information is complete.)

TABLE D.1.

SAGINAW DAY SCHOOL ELEMENTARY OFFICIAL MEMBERSHIP - SEPTEMBER 26, 1980

	TOTAL	PRE-K	KIN.	1	2	3	4	5	6	SPL.	TOTAL MEMBERSHIP
E. Baillie	377	32	35	46	54	27	61	63	36	23	345
Coulter	314	37	43	31	37	47	40	42	29	8	277
Emerson	593	36	82	82	94	82	78	60	60	19	557
Fuerbringer	324	-	50	42	39	50	45	35	31	32	324
N. Haley	459	63	43	63	57	55	47	47	43	41+	396
Handley	247	-	67	25	24	20	19	25	36	31	247
Heavenrich	517	-	79	76	73	92	64	68	65	-	517
Herig	399	-	66	52	55	51	63	63	49	-	399
Houghton	494	39	62	68	57	52	74	60	48	34	455
Jerome	275	-	45	46	48	41	31	31	33	-	275
Jones	433	29	71	77	59	61	40	53	43	-	404
Kempton	167	-	14	18	23	27	26	23	27	9	167
Longfellow	637	40	91	93	98	92	86	64	73	-	597
Longstreet	230	-	35	20	27	23	20	28	22	55*	230
J. Loomis	520	38	71	72	87	61	64	60	67	-	482
Merrill Park	411	-	47	71	60	54	70	48	52	9+	411
C. Miller	220	-	20	28	30	27	39	31	29	16	220
J. Moore	277	-	35	45	37	42	37	30	35	16	277
Morley	346	38	42	35	41	47	43	40	51	9+	308
Potter	306	13	38	47	41	48	42	27	36	14	293
J. Rouse	437	37	67	67	62	64	48	46	46	-	400
Salina	265	20	26	45	32	35	29	33	28	17	245
South Elem.	252	-	33	32	25	44	38	41	39	-	252
Stone	346	-	41	52	44	49	47	35	50	28	346
Webber Elem.	542	-	78	68	81	90	86	72	67	-	542
Wilmarth	376	-	41	47	56	47	60	52	73	-	376
TOTAL	9,704	422	1,322	1,348	1,341	1,328	1,297	1,177	1,168	361	9,342

8 Students are co-enrolled in Kindergarten and Early Childhood Program.

* General Education Total includes Special Education Resource Room students as follows:

Belle Haley	13
Merrill Park	10
Morley	10

Special Education Total does not include these students.

11-19-80

1980-81

TABLE D.2.

CITY OF SAGINAW SCHOOL DISTRICT RACIAL & ETHNIC COUNT

SCHOOL	AMERICAN INDIAN	BLACK	ASIAN AMERICAN	LATINO OR HISPANIC	WHITE	TOTAL
Edith Baillie	3	370	-	3	1	377
Coulter	7	242	4	36	25	314
Emerson	-	502	-	74	18	594
Fuerbringer	5	17	1	26	275	324
M. Haley	2	263	1	104	89	459
Handley	3	20	-	10	214	247
Heavenrich	10	481	-	20	6	517
Herig	-	3	1	46	349	399
Houghton	-	445	-	45	5	495
Jerome	11	7	-	46	211	275
Jones	-	390	-	34	4	434
Kempton	2	34	3	15	113	167
Longfellow	-	409	2	102	64	637
Longstreet	-	174	1	27	28	230
M. Loomis	2	392	-	96	30	520
Merrill Park	5	13	-	62	331	411
C. Miller	9	4	1	8	198	220
J. Moore	6	9	1	31	230	277
Morley	-	332	-	13	1	346
Potter	2	286	-	18	1	307
J. Rouse	3	153	-	193	88	437
Salina	-	235	13	12	4	264
South Elem.	-	34	2	53	163	252
Stone	7	31	11	52	245	346
Webber Elem.	10	260	-	153	119	542
Zilwaukee	6	-	3	14	353	376
Total	93	5172	44	1292	3166	9767
Arthur Hill	54	249	9	160	1308	1780
Saginaw High	9	1334	-	101	79	1523
Total	63	1583	9	261	1387	3303
Arthur Eddy	-	397	-	12	-	409
Central Junior	3	609	-	35	21	668
North Intermed.	31	222	2	90	577	922
South Intermed.	35	85	2	91	583	796
Webber Junior	9	529	6	158	128	830
Total	78	1842	10	386	1309	3625
Continuation	-	70	-	5	4	79
Millet Center	2	93	1	25	193	314
Holland Ed. Cent.	-	13	-	3	9	25
Total	2	176	1	33	206	418
Grand Total	236	8773	64	1972	6068	17,113

TABLE D.3.

SAGINAW DAY SCHOOL OFFICIAL MEMBERSHIP - SEPTEMBER 26, 1980

SUMMARY

High School Total	3,368
Junior High Total	3,639
Elementary Total	9,342
Millet Learning Center	310
Holland Educational Center	22
	16,681

Pre-Kindergarten Program - Federally Funded

422 Students

Special Education is figured on Full Time Equivalency and the head count will not coincide with the Official Membership Report.

PRE-KINDERGARTEN

Baillie	32
Coulter	37
Emerson	36
H. Haley	63
Houghton	39
Jones	29
Longfellow	40
J. Loomis	38
Morley	38
Potter	13
J. Rouse	37
Salina	20
	422

SPECIAL EDUCATION ELEMENTARY

Baillie	23
Coulter	8
Emerson	19
Fuerbringer	32
+H. Haley	41
Handley	31
Houghton	34
Kempton	9
*Longstreet	55
+Merrill Park	9
Miller	16
J. Moore	16
+Morley	9
Potter	14
Salina	17
Stone	28
	361

SPECIAL EDUCATION SECONDARY

Arthur Hill	120
Saginaw High	139
Arthur Eddy	44
Central Junior	60
North Intermediate	56
South Intermediate	50
Webber Junior	57
	526

* 8 Students are co-enrolled in Kindergarten and Early Childhood Program.

+ General Education Total includes Special Education Resource Room students as follows: Helle Haley 13
Merrill Park 10
Morley 10

Special Education Total does not include these students.

11-19-80

APPENDIX E

TITLE I PREKINDERGARTEN ACTIVITY OBSERVATION CHECKLIST

1979-80

Teacher's Name _____ Observer's Name _____

School _____ Date _____ Length of Observation _____

Number preschoolers _____

Product Objective Referent	Type of Activity*	Check if Activity Occurred During Observation Period	
		(✓)	Example
	Behavior Modification Techniques**		
	Interest Centers**		
13	Gross Motor Coordination**		
10	Eye-Hand Coordination (Gross and Fine Motor and Manipulative)**		
1	Properties of Objects; i.e., shape, color, hardness (five senses)		

*Refer to ESEA Title I Pre-School Examples of Pre-School Activities Sheet for a detailed explanation of the types of activities.

**These activities plus some aspect of work on physical knowledge should be part of the daily classroom activity.

Product Objective Referent	Type of Activity*	Check if Activity Occurred During Observation Period	
		(√)	Example
	Rules of Behavior (school and traffic)		
2	Social Knowledge (i.e., work roles)**		
3	Grouping and Regrouping (i.e., classification)		
4	Transitive Relations (i.e., length, height, weight, shades, hardness)		
	Conservation of Numbers by One-to-One Comparison (i.e., matching, pouring, getting coats, rearranging collections)		
	Topological Space (i.e., in-out, under, etc.)		
	Linear Order (i.e., straight lines, counting)		
	Parts of a Whole		
12	Copying Specific Shapes (i.e., cutting, pantomine, drawing)		

*Refer to ESEA Title I Pre-School Examples of Pre-School Activities Sheet for a detailed explanation of the types of activities.

**These activities plus some aspect of work on <u>physical knowledge</u> should be part of the <u>daily</u> classroom activity.

Product Objective Referent	Type of Activity*	Check if Activity Occurred During Observation Period	
		(√)	Example
11	Temporal Ordering of Events		
	Use of Body to Represent an Object		
	Use of Objects to Represent Other Objects		
	Human Utterance Which Represent Objects		
	Use of Three-Dimensional Model to Represent Object		
	Two-Dimensional Representation of Objects in Drawing		
14-16	Record of Parental Participation Being Maintained		

*Refer to ESEA Title I Pre-School Examples of Pre-School Activities Sheet for a detailed explanation of the types of activities.

(Key for Classroom Activity Observation Checklist)

ESEA TITLE I - PRE-SCHOOL

Examples of Pre-School Activities According to Product and Process Objectives

Type of Activity	Activity Examples	
Behavior Modification	-One-to-one relationship with an adult -Seeking adults as resources -Consistent classroom environment--inner control--freedom and responsibility -Sharing, selecting partners, initiating activities with others -Positive reinforcement -Continue with a task -Exhibit curiosity, questions, explore, experiment -Creativity--different ways to approach a task	
Gross Motor Coordination (large body movements, climbing, walking, rolling)	-Rhythms -Dancing -Jungle Gym -Free play activities -Balance Beam -Mats - tumbling -Play All Equipment -Jumpin Jiminy -Jump Roles - forming circles with activities -Jumping Jacks -Duck Duck Goose -Squirrel in Tree	-Johnny Works With One Hammer -Bear Hunt -Acting out Mother Goose Rhyme -Rhythm Estamae -Dodge Ball -Balls & Skate Board -Play House -Roller Skates -Snowman Activities -Up the Steps
Fine Motor Activities--Eye-Hand Coordination (use of classroom tools and materials--cutting, pasting, tearing)	-Art work -Writing on the board -Finger painting -Folding -Stirring pudding -Peg boards -Pouring -Geoboards -Puzzles -Cuisenairre Rods -Sorting beads and buttons -TRY -Building Blocks -Lacing	-Weaving -Chalk Boards -Flannel Boards -Clay -Sand box -Water play -Spreading peanut butter -Coats - buttons and zippers -Clean up time -Finger plays -Using musical instruments

(continue)

(ESEA TITLE I - PRE-SCHOOL continued)

Type of Activity	Activity Examples	
Properties of and Appropriate Behavior for Exploring Properties of an Object (shape, color, hardness--using the five senses. Changing shades, measuring, weighing)	-Making apple sauce, soups, cookies, etc. -Smelling and handling fruits and vegetables -Sawing wood -Tinkertoys -Sand paper activities -Feeling activities -Snacks - (mixtures) -Snow experiments -Bubble blowing -Straw painting -Furry and other textured toys -Fast and slow inclined plane -Paper mache	-Growing plants from seed -Cutting -Freezing -Heating -Rolling -Twisting -Frosting -Jello -Butter -Cakes -Paint mixing -Sinking and floating -Color macaroni -Play dough
Knowledge of Rules that Apply to Learners, i.e., health, and safety	-Stories -Visitors - nurse -Health Lessons	-Safety Lessons -Line-up activities -Taking turns
Social Knowledge (world of work and roles of workers)	-Books -Field Trips -Films -Visitors -Role-playing -Helpers in the room	-Community workers -School workers -Visiting patrolmen -Postman -Kandy Keg Field Trip
One Criterion Classification, Shifting to a second criterion among an array of objects (grouping shifting from one criterion to another)	-Color - blocks -Shape -Size -Texture -Tone -Utility -Smell -Taste -Calendar	-Sorting -Attendance - number of girls -Attendance - number of boys -Putting toys away -Doll House -Doll Dishes

(continue)

(ESEA TITLE I - PRE-SCHOOL continued)

Type of Activity	Activity Examples	
Relations Among Transitive Relationships (seriation--comparing and arranging things according to a given dimension by transitive relations)	-Length -Height -Weight -Shades of color -Hardness -Softness -Cuisennaire rods -Block tower building -Texture activities	
Convervation of Number by One-to-One Comparison (gross comparison between collections; comparisons by one-to-one correspondence)	-Collections - rearrangement of -Lunch activities -Setting table -Matching -Calendar -Passing anything -Weather	-Getting coats -Right boot -Pouring activities
Topological Space (meaning of in-out, over-under, in front of, in back of, in relation to self, toys, pictures)	-Ring Around the Rosies -Squirrel in a Tree -Hokey Pokey -When You're Up Your Up -In and Out the Windows -Bear Hunt	-Directions from adult -Three Billy Goats Gruff -Jack in the Box -London Bridge -Popcorn -Use of box
Topological Relationships Concerning Linear Order (structure of space)	-Games - straight line -Role-playing -Manipulation of Object (rods, blocks, toys) -Poetry -Prose	-Counting days till -Finger plays -Bear Hunt -AAA -Ten Little Indians
Parts of a Whole	-Clay -Paper cutting -Puzzles - inlaid -Coloring -Sawing -Cooking -Body Parts -Growing plants -Filling partial measurements	-Gingerbread Man -Bread and rolls -Strawberry jam -Ripping cloth strips -Torn paper artwork -Loops - artwork -Placement of toys in specific unit or box

(ESEA TITLE I - PRE-SCHOOL continued)

Type of Activity	Activity Samples	
Copying of Specific Shapes	-Line drawings -Sand drawing -Paper cutting -Cookie cutting with clay -"Simon Says" -Tracing -Rubbing -Pegboard	-Geoboards -TRY -Writing chalkboard -Directed copying activity -Pantomine -Exercises
Temporal Ordering of Three or Four Events (structuring time)	-Show and Tell -Story - book -Role-playing -Science experiments -Calendar -Preparation art, lunch, cleanup home bound	-Growth stages -Finger play -Farmer in Dell -Audio-visual materials
Use of Body to Represent Objects (i.e., pretending to hold a telephone receiver)	-Musical games -Role-playing -Shadow Plays -Dramatization -Elephant Walk	-Seals -Bunny Hop -Exercise Records -Bodies for race cars
Use of Objects to Represent Other Objects (props) (representation at symbol level)	-Story telling by using props -Role-playing -Puppets -Flannel Board	
Uttering of Sounds to Represent Objects	-Role-playing -Musical game -Puppets	-Flannel board -Dramatization
Use of Three-Dimensional Models to Represent Objects (representation at symbol level)	-Pretend telephone -Playing house -Using blocks -Buggy for cars	-Educubes for drums -Domino blocks for walkie talkie
Two-Dimensional Representation of Objects in Drawings (eye-hand manipulation)	-Read aloud of picture story book -Manipulation of packets of cards which contain pictures of common objects -Drawing on chalkboard, easel, coloring paper -Play with water and brush, pitcher, or sprinkling can -Play teacher	

APPENDIX F

TABLE F.1.
SUMMARY BY OBJECTIVE OF 1978-79 TITLE I
PREKINDERGARTEN MASTERY CRITERIA

Objective Number	Component	Prekindergarten S.O.R.T. Test Items		Mastery Criteria
1	Cognitive	1, 2, 3	80%	2 of 3
2	Cognitive	4, 5, 6, 7	80%	3 of 4
3	Cognitive	8, 9	50%	2 of 2
4	Cognitive	10, 11	70%	1 of 2
5	Cognitive	12, 13	50%	1 of 2
6	Cognitive	14	85%	1 of 1
7	Cognitive	15	80%	1 of 1
8	Cognitive	16	65%	1 of 1
9	Cognitive	17	50%	1 of 1
10	Psychomotor	41, 42, 43, 44	80%	3 of 4
11	Psychomotor	45, 46	65%	2 of 2
12	Psychomotor	47, 48, 49, 50	65%	3 of 4
13	Psychomotor	51, 52, 53, 54	80%	3 of 4
14	Parent Participation		60%	5 times
15	Parent Education Program		60%	3 times
16	Parent Education Program		80%	9 activities

1978-1979 TITLE I PREKINDERGARTEN PRODUCT OBJECTIVES

1. Physical Knowledge

   80% of the pupils will correctly respond to at least two of three items related to objectives 1.

2. Social Knowledge (social roles)

   80% of the pupils will demonstrate knowledge of social roles by correctly responding to at least three of four items related to objective 2.

3. Knowledge: Classification

   50% of the pupils will successfully apply two criteria for sorting: color and/or form.

4. Knowledge: Logical-Mathematical-Seriation

   70% of the pupils will achieve mastery of at least one of two items related to objective 4.

5. Spatiotemporal Knowledge: Structuring of Time

   50% of the pupils will respond correctly to at least 50% of items related to objective 5.

6. Expressive Language: Labeling

   85% of the pupils will label at least four objects in the birthday party picture.

7. Expressive Language: Syntax

   80% of the pupils will use sentences of at least five words to describe the birthday party picture.

8. Expressive Language: Fluency

   65% of the pupils will use at least three of five elements of fluency in their description of the birthday party picture.

9. Expressive Language: Plot Extension

   50% of the pupils will use at least one element of plot extension in their description of the birthday party picture.

10. Fine Motor Coordination

    80% of the pupils will complete successfully at least three of four of items related to objective 10.

11. Spatiotemporal Knowledge: Structuring of Space (order)

    65% of the pupils will correctly respond to topological relationships of order or pattern.

12. Representation at the Symbol Level: Specific Shapes

    65% of the pupils will copy successfully three of four shapes.

13. Gross Motor Coordination

    80% of the pupils will successfully complete at least three of four items related to objective 13.

14. Parent Participation

    60% of the adult members of the prekindergarten family will participate in school activities at least five times per year.

15. Parent Education Program: Friday Meetings

    60% of the adult members of the prekindergarten family will participate in at least three friday meetings.

16. Parent Education Program: Home Work Activities

    80% of the adult members of the prekindergarten family will complete at least nine prekindergarten home activities and return them to school.

TABLE G.1.
NORMED DATA REFERENCES FOR BEHAVIORS INCLUDED ON THE PREKINDERGARTEN READINESS SCREENING DEVICE

Test Behavior	Norm in Months	Reference	Comment
1. Says last name	30-36	Frankenburg and Dodds, 1969; Gesell, 1940; Sheridan, 1968	
	36-42	Doll, 1966; Doll, 1965	
	60-72	Gesell Human Development Institute, 1979	Specific behavior normed: gives first and last name (interview)
	73% of 48-59 80% of 60-71 83% of 72-78	Caldwell, 1967	Normed on disadvantaged population
2. Says age	36-41 (girls)	Gesell Human Development Institute, 1979	
	48-53 (boys)	Gesell Human Development Institute, 1979	
	60-72	Doll, 1966; Sheridan, 1968	
	76% of 48-59 82% of 60-71 82% of 72-78	Caldwell, 1967	Normed on disadvantaged population
3. Points to neck	36-47	Danger, Lyons, and Gerber, 1972	
	88% of 48-59 94% of 60-71 95% of 72-78	Caldwell, 1967	Normed on disadvantaged population
4. Identifies elbow	68% of 48-59 76% of 60-71 80% of 72-78	Caldwell, 1967	Normed on disadvantaged population
5. Gives use of ear	60-71	Danger, Lyons, and Gerber, 1972	

Table G.1. (Cont'd.)

Test Behavior	Norm in Months	Reference	Comment
6. Matches same colored object (blue)	30-35	Zehrback, 1975	Specific behavior normed - matches block of four colors (3 of 4 correct)
	30-36	Stutsman, 1963	Specific behavior normed - matches four colors
	36-42	Gesell, 1940	Specific behavior normed - matches 2-3, confuses blue and green
	36-42	Slosson, 1964	Specific behavior normed - matches 3 colors
7. Picks up object of the color named (green)		Danger, Lyons, and Gerber, 1972	
8. Names colored object (yellow)	36-42	Frankenburg and Dodds, 1969	Specific behavior normed - identifies 3 of 4 colors
	45	Quick, Little, and Campbell, 1974	Specific behavior normed - identifies three colors: "Show me the blue one," etc.
	48-54	Zimmerman, Steiner, and Evatt, 1969	Specific behavior normed - identifies 1 of 4 colors
	48-59	Danger, Lyons, and Gerber, 1972	
	60-66	Banham, 1963	Specific behavior normed - labels 4 colors
	66-72	Doll, 1966	Specific behavior normed - labels 4 colors
9. Names colored object (black)	60-71	Danger, Lyons, and Gerber, 1972	
	60	Quick, Little, and Campbell, 1974	Specific behavior normed - names or labels at least six colors

Table G.1. (Cont'd.)

Test Behavior	Norm in Months	Reference	Comment
10. Counts to five	48-59	Danger, Lyons, and Gerber, 1972 Supplemental references to other rote counting	
	30-36	Doll, 1966 and 36-42 Doll, 1965	Specific behavior normed - counts to 2
	42-48	Doll, 1966 and 48-54 Gesell, 1940	Specific behavior normed - counts to 3
	54-60	Gesell, 1940	Specific behavior normed - counts to 4
11. Hands four blocks from a pile of five	48-54	Doll, 1966 and Gesell, 1940	
	48-59	Danger, Lyons, and Gerber, 1972	
12. Says number of blocks (three)	54	Quick, Little, and Campbell,	Specific behavior normed - counts "three" meaningful, i.e., "give me three cars."
	48-59	Danger, Lyons, and Gerber, 1972	
13. Picks up all the blocks	48-59	Danger, Lyons, and Gerber, 1972	
	58% of 48-59 73% of 60-71 85% of 72-78	Caldwell, 1967	Specific behavior of norm - gives everything (3 cars plus box) to examiner and normed on disadvantaged population
14. Walks backwards toe to heel two or more steps	54-59	Zehrbach, 1975	Specific behavior normed - walks backwards toe to heel for three steps (toe and heel no more than one inch apart)
	56	Frankenburg and Dodds, 1969	
15. Carries out 3 part command	54	Cattell, 1950	Score correct if separate in proper order
	57	Quick, Little, and Campbell, 1974	Specific behavior normed - follows three commands in correct sequence

Table G.1. (Cont'd.)

Test Behavior	Norm in Months	Reference	Comment
16. Gives use of books	48-59	Danger, Lyons, and Gerber, 1972	
17. Draws cross given model (copying without demonstration)	36-42	Goodenough, Maurer, and Wagenen, 1940	
	42-48	Stutsman, 1963; Frankenburg and Dodds, 1969; Gesell Human Development Institute, 1979	
	44 (boys)	Beery and Buktenica, 1967	
	48-53	Zehrback, 1975	
	48-54	Danger, Lyons, and Gerber, 1972	
	49 (girls)	Beery and Buktenica, 1967	
	51	Quick, Little, and Campbell, 1974	
18. Hops on one foot for two or more hops	30-36	Bayley, 1969	
	36-48	Frankenburg and Dodds, 1969 and Sheridan, 1968	
	42-47 (girls)	Gesell Institute of Human Development, 1979	
	48-53	Zehrback, 1975	
	51	Quick, Little, and Campbell, 1974	
	54-59 (boys)	Gesell Institute of Human Development, 1979	
19. Throws ball five feet within reach of teacher	36-42	Doll, 1966	Specific behavior normed - underhand throw
	36-48	Gesell, 1940	Specific behavior normed - overhand throw
	48-54	Banham, 1963.	Specific behavior normed - overhand throw
20. Points to shape like a wheel	76% of 48-59 94% of 60-71 95% of 72-78	Caldwell, 1967	Normed on disadvantaged population

Table G.1. (Cont'd.)

Test Behavior	Norm in Months	Reference	Comment
21. Points to shape like a tent	50% of 48-59 79% of 60-71 73% of 72-78	Caldwell, 1967	Normed on disadvantaged population
22. Points to shape like a stick	71% of 48-59 84% of 60-71 89% of 72-78	Caldwell, 1967	Normed on disadvantaged population
23. Tells which is bigger, a tree or a flower	30-36	Doll, 1966	
	42-48	Cattell, P., 1960	
	45	Quick, Little, and Campbell, 1974	Specific behavior normed - understands "bigger" and "longer" concept
	68% of 48-59 85% of 60-71 94% of 72-78	Caldwell, 1967	Normed on disadvantaged population
	64	Slosson, 1964	
	60-71	Danger, Lyons, and Gerber, 1972	
24. Traces around diamond template	88% of 48-59 90% of 60-71 87% of 72-78	Caldwell, 1967	
	42	Gesell, 1940	
	45	Quick, Little, and Campbell, 1974	
25. Tells a nursery rhyme or song	36-48	Gesell, 1940; Doll, 1965; Sheridan, 1968	
	45	Quick, Little, and Campbell, 1974	

Table G.1. (Cont'd.)

Test Behavior	Norm in Months	Reference	Comment
26. Says "yellow jello"	51	Quick, Little, and Campbell, 1974	Specific behavior normed - mastery of Y, J, K, G, F, and D sounds
	54	Gesell, 1940	Specific behavior normed - articulation not in infantile style
27. Engages in conversation-like speech	42	Quick, Little, and Campbell, 1974	Specific behavior normed - complete sentences used consistently
	60-72	Sheridan, 1968	Specific behavior normed - speaks fluently and correctly except for confusion of s, f, th

SAGINAW OBJECTIVE REFERENCED TESTS

PREKINDERGARTEN

Rev. 1978

par

is

wro

are

sho

tha

for

and

to

te

et

th

pu

to

ap

o

y

DIRECTIONS FOR ADMINISTERING PREKINDERGARTEN S.O.R.T.

- Read these thoroughly before testing -

All directions that are for teacher use only will be in parentheses ( ) and are not to be read to the pupil. This test is to be administered on a one-to-one basis.

Even though the test is immediately scored on a right or wrong basis by the person administering it, all pupil responses are to be accepted, i.e., the content of a pupil's response should be accepted by the test administrator.

The person administering the test should, however, make sure that the pupil clearly understands the task that s/he is to perform, i.e., s/he should be told how to respond to the questions and this may involve correcting the pupil.

A separate machine readable answer sheet will be provided to those administering the test and it should be marked as the test is administered. Use these symbols for scoring:

A for correct responses.
B for incorrect or no responses.

Print the pupil's name, last name first, then first name, etc., in the boxes provided on the scoring sheet. Then blacken the letter box below the letter which matches those in the pupil's name. Also print and darken in the boxes corresponding to the birth date (month and year) and sex of each pupil in the appropriate boxes provided on the answer sheet. Print the name of the school and your name (instructor) on the lines provided. Make every attempt to administer all items to pupils being tested.

Take as much time as you think is necessary to administer the test to each pupil. This may involve administering the test in a number of sittings.

Since you are to record correct and incorrect responses, your directions for each test item will also have a section describing what the correct or acceptable responses are so that you can appropriately mark your scoring sheet.

Familiarize yourself with all testing materials before administering the test. Make sure that you have a complete set of manipulative materials and flash cards before beginning the test. In addition, read all the specific directions for administering the test items before testing begins. This practice should help you more quickly score each response since you will be somewhat knowledgeable of the acceptable response charts provided in the test instructions.

PREKINDERGARTEN S.O.R.T. TEST ITEMS

- I. COGNITIVE DEVELOPMENT SUBTEST -

Remember, all statements in parentheses ( ) are intended for your use and are not to be read to the child. In recording answers on your answer sheet, you should code A - correct response and B - incorrect or no response.

1. SAY, "Let's play a game where you have to tell me about things you cannot see."

(Hand the pupil feely sock box Number 1. It contains a metal zipper.)

SAY, "Put your hand in the sock. Keep your hand in the sock. Take the thing that is in the sock and hold on to it. Tell me about it."

(Pause. Listen for one of these acceptable responses and mark your scoring sheet accordingly.)

Acceptable Responses

-- name of the object
-- shape of the object
-- use of the object
-- name of the material of the object
-- texture of the object

SAY, "Let's take a look at it. Now, let's do another one."
(Put away box Number 1 making sure that the zipper is put back, take out box Number 2.)

2. (Hand the pupil feely sock box Number 2. It contains a toothbrush.)

SAY, "Put your hand in the sock. Keep your hand in the sock. Take the thing that is in the sock and hold on to it. Tell me about it."

(Pause. Listen for one of these acceptable responses and mark your scoring sheet accordingly.)

Acceptable Responses

-- name of object
-- shape of the object
-- use of the object
-- name of the material of the object
-- texture of the object

SAY, "Let's take a look at it. Now, let's do another one."
(Put away box Number 2 making sure that the tooth-brush is put back, take out box Number 3.)

(Hand the pupil feely sock box Number 3. It contains a plastic egg.)

3. SAY, "Put your hand in the sock. Keep your hand in the sock. Take the thing that is in the sock and hold on to it. Tell me about it."

(Pause. Listen for one of these acceptable responses and mark your scoring sheet accordingly.)

Acceptable Responses

-- name of the object
-- shape of the object
-- use of the object
-- name of the material of the object
-- texture of the object

SAY, "Let's take a look at it."

(Put away box Number 3 making sure the egg is put back.)

4. SAY, "Now let's take a look at some pictures and talk about them."

(Show the child the picture marked with the Number 4 on the back. As you are holding it follow these directions.)

SAY, "Tell me who this worker is. What does s/he do?"

(Pause for response, listening for one of the acceptable responses listed below.)

Acceptable Responses

-- name of the role or title of the worker

OR

-- a description of what s/he does or how the worker helps us.

(Mark your scoring sheet accordingly.)

5-7. (Put away picture Number 4 and continue following the same directions for pictures 5, 6, and 7. Remember to mark on your scoring sheet after each question.)

---

8. SAY, "In just a minute we will play a game with some candies which should be lots of fun."

(Open the envelope marked item Number 8 and randomly place candies that it contains in front of the pupil.)

SAY, "Some candies are yellow, some are green, some are round, and some are long. Put the candies that are alike into two piles."

(Pause for the child to group the candies. Make sure that one of the groups is correct according to the acceptable responses listed below.)

Acceptable Responses

-- grouping according to color
-- grouping according to form

(Mark your scoring sheet accordingly.)

9. SAY, "In just a minute we will play a game with some circles and triangles which should be lots of fun."

(Open the envelope marked item Number 9 and randomly place the shapes in front of the pupil.)

SAY, "Put the shapes that are the same into two piles."

(Pause for the child to group the shapes Make sure that the groups are correct according to the acceptable response listed below.)

Acceptable Response

-- grouping according to form

(Mark your scoring sheet accordingly.)

---

10. SAY, "Now let's play with some toy people. They are a family."

(Remove toy dolls from envelope marked item No. 10 and allow child to play with and talk about father, mother, sister or brother and baby toys or their own family members.)

SAY, "Now, can you put this family from the biggest to the smallest?"

(Pause for the child to arrange the dolls from biggest to smallest or the reverse order. Make sure that the arrangement is correct according to the acceptable responses listed below.)

Acceptable Responses

-- all four dolls from biggest to smallest

OR

-- all four dolls from smallest to biggest

(Mark your scoring sheet accordingly.)

11. SAY, "Now let's take a look at some pictures and put them in order."

(Open the envelope marked item Number 11 and randomly place the four pictures in front of the pupil.)

SAY, "Here are four girls. Some of the girls are tall, some are short. Put the girls in a row from tallest to shortest."

(Provide a ruler as base. Pause for the child to arrange the girls. Make sure that the arrangement is correct according to the acceptable responses listed below.)

Acceptable Responses

-- all four pictures from tallest to shortest

OR

-- all four pictures from shortest to tallest

(Mark your scoring sheet accordingly.)

SAY, "Let's play a game with some pictures and stories. I will read you a story. Then you will make the pictures tell me what happened. You will give me the picture that happened first, next, and last."

(Open envelope marked sample, 12 and 13. Take out pictures for the sample item.)

SAMPLE EXERCISE

SAY, "Let's do the first one together. Listen to this story. 'Mary is riding her bicycle to school. She locked it up. Then she played ball with the kids.' Now let's put the pictures together so they tell the same story." (Teacher hands the pictures to the child.) "Give me the picture that happened first."

(Pause for answer and correct child if s/he has not understood directions.)

SAY, "Give me the picture that happened next."

(Pause for answer and correct child if s/he has not understood directions.)

SAY, "Give me the picture that happened last."

(If child gives incorrect sequence, teacher tells the story and presents the pictures in the correct order.)

(Take out pictures marked Number 12.)

12. SAY, "Let's do another picture story. Listen to this story. 'Danny broke a glass while washing dishes. He swept up the glass. He put the broken glass in the trash can.' (Teacher hands the pictures to child.) "Give me the picture that happened first." (Pause for the correct picture.) "What happened next?" (Puase for the correct picture.) "What happened last?"

(Pause for the child to answer the questions. Make sure that the answer is listed below as an acceptable response.)

Acceptable Response

-- all three pictures in correct time order sequence even if backwards.

(Mark your scoring sheet accordingly and put the pictures away.)

(Take out pictures marked Number 13.)

13. SAY, "Let's do another picture story. Listen to this story. 'Jane and her two friends climbed the tree. The branch Jane was on broke and she fell. Jane broke her leg and had to walk on crutches.' Now put the pictures together so they tell the same story." (Teacher hands the three pictures to child.) "Give me the picture that happened first." (Pause for the picture.) "What happened next?" (Pause for the picture.) "What happened last?"

(Pause for the child to answer the questions. Make sure that the answer is listed below as an acceptable response.)

Acceptable Response

-- all three pictures in correct time order sequence even if backwards

(Mark your scoring sheet accordingly and put the pictures away.)

14. SAY, "I have a picture here out of a story book. It's part of the story, but the words are missing. Would you look at my picture and help me with the story? "

(Teacher hands child the picture from folder marked Number 14.)

SAY, "Tell me what you see in this picture."

(Pause for the child to answer. Make sure that the answer is listed below as an acceptable response.)

Acceptable Responses

— name at least four objects in picture (need not identify correctly)

For example: animals
balloons
candles
koolaid
juice
chairs
table
hats
cake

Incorrect Responses

— did not talk
— named less than four objects
— gave irrelevant responses

(Mark your scoring sheet accordingly.)

(Child continues to use picture marked Number 14.)

15. "Tell me what you think is happening in the picture?

(Pause for the child to tell the story. Make sure that the answer is listed below as an acceptable response.)

Acceptable Response

— uses a sentence of 5 or more words

Incorrect Responses

— child does not talk
— uses sentences of four words or less
— uses phrases

(Mark your scoring sheet accordingly.)

16. (Score story given for item 15 in terms of acceptable responses given below.)

Acceptable Response

-- uses at least 3 of 5 of the listed elements of fluency. *

Incorrect Response

-- uses less than 3 of the listed elements of fluency. *

* Fluency consists of additional responses using:
  -- modifiers (uses adjectives or adverbs.)
  -- spatial elements (uses prepositions indicating position.)
  -- number words
  -- emotional or feeling words
  -- sequence (uses phrases to describe a series of events.)

(Mark your scoring sheet accordingly.)

17. (Child continues to hold the picture from the folder marked Number 15.)

SAY, "What do you think will happen next? What will they do when the party is over?"

(Pause for the child to answer. Make sure that the answer is listed below as an acceptable response.)

Acceptable Response

-- child uses 1 or more of the elements listed below as plot extension. *

*Plot extension consists of:
  -- inferences
  -- predictions
  -- cause and effect
  -- conclusions

Incorrect Response

-- child does not use plot extension. *

* Plot extension consists of:

  -- inferences
  -- predictions
  -- cause and effect
  -- conclusions

(Mark your scoring sheet accordingly and put the picture away.)

- II. PSYCHOMOTOR ABILITIES SUBTEST -

41. (From envelope marked Number 41, ask pupil to fold a 5" x 5" sheet of paper in half. Teacher demonstrates with a sample.)

SAY, "Fold the paper in half."

Acceptable Response

-- using ruler, folds should show an accuracy ± 3/8" in any direction.

(Mark scoring sheet accordingly.)

---

42. (Using the same folded sheet, ask pupil to open the sheet and cut the paper on the fold.)

SAY, "Now open the sheet and cut the paper on the fold line."

(Teacher demonstrates with his/her sample.)

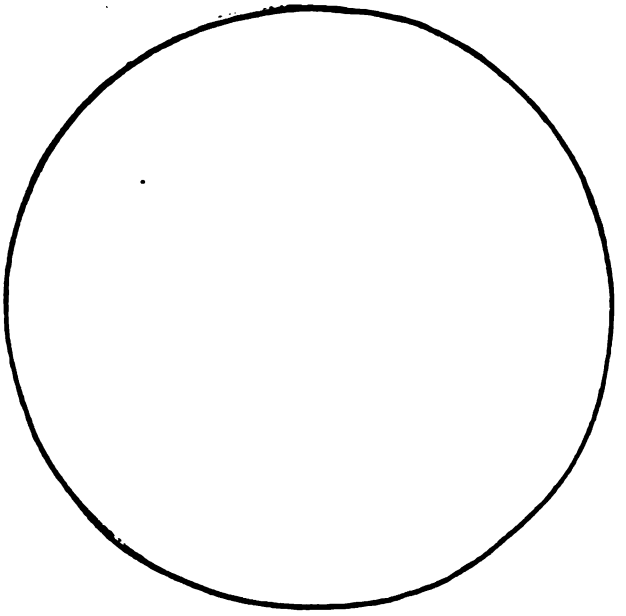
Acceptable Response

-- using ruler, cuts should be ± ½" from the fold.

(Mark scoring sheet accordingly.)

43. (Using a crayon from envelope Number 43, ask pupil to color inside the outline of the circle.)

SAY, "Color inside this circle."



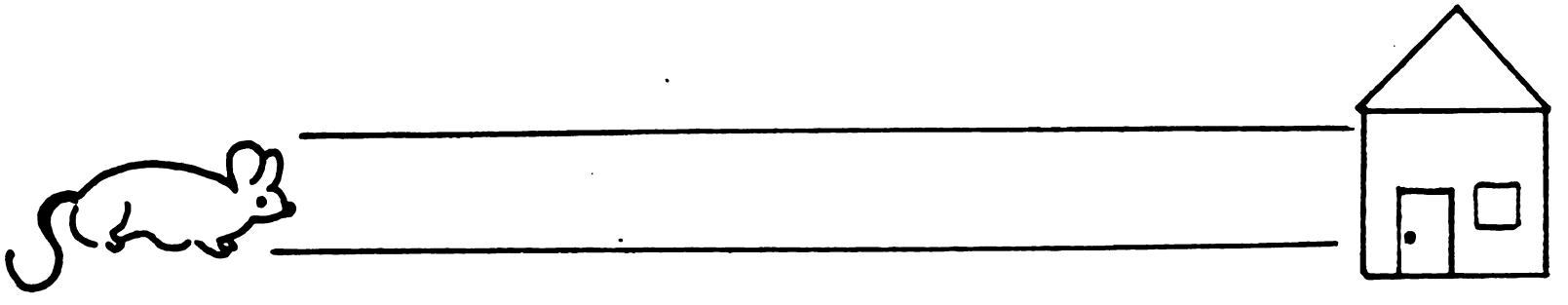
Acceptable Response

-- using ruler, coloring marks should not exceed 1/2" at any point and approximately 2/3 rds of circle should be colored.

(Mark scoring sheet accordingly.)

44. (Using a crayon from envelope Number 44, ask pupil to draw a line between the two lines.)

SAY, "Draw a line between the two lines from the mouse to the house."



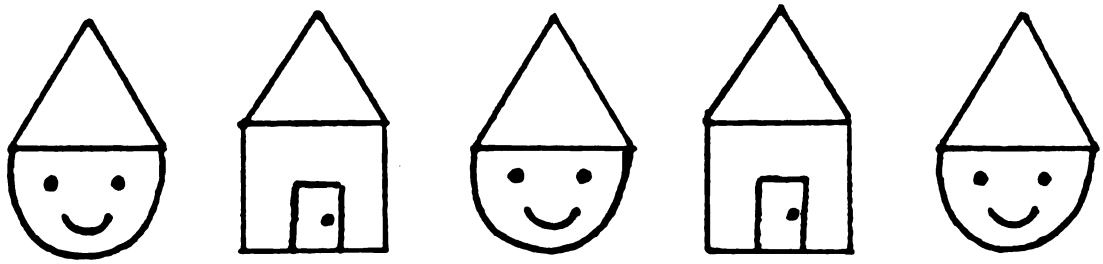
Acceptable Response

-- Crayon line must be within parallel lines and connect the mouse to the house or come within at least 1/2" of touching both the mouse and the house.

(Mark scoring sheet accordingly.)

45. (Using cut-out forms from envelope Number 45, place them on the table facing the child in the order shown below. Then take a similar set from envelope Number 45 and ask the child to make the same pattern.)

SAY, "You make your row look just like mine."



Acceptable Response

-- Linear order must be the same as the example.

(Mark scoring sheet accordingly.)

(Child must be seated across from the teacher. Teacher places his/her 5 toy cars from envelope Number 46 on the oaktag circle. The teacher puts down the parking strip, one in front of the child and another at least 10 inches away from that one and parallel to it.)

SAY, "We are going to build parking lots. First, you watch how I park my cars and trucks."

(From the circle the teacher takes 5 cars and places them on teacher parking strip #2 in the same predetermined order for all children as printed on parking strip. Teacher then places the child's 5 toy cars on the oaktag circle and asks the child to park his/her cars on child's parking strip #1 to look just like the teacher's.)

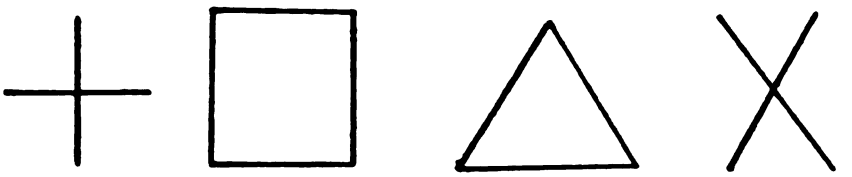
SAY, "Park your cars just like mine."

Acceptable Response

-- Linear order of cars must be the same as the teacher's order according to color.

(Mark scoring sheet accordingly.)

47-50 (Using cards from envelope Number 47-50, show one card at a time in the following order. Hand an extra sheet of paper to the child to draw the figures.)



SAY, "Draw a shape like this one."

Acceptable Response

-- See Appendix A for acceptable drawings as shown in Administration and Scoring Manual for the Developmental Test of Visual Motor Integration.

---

51. (Given the directive (opportunity) to hop on one foot the child will be able to take five consecutive hops on either foot.)

Acceptable Response

-- Successful performance of the above activity.

(Mark scoring sheet accordingly.)

---

52. (Given a mark on the floor the child will be able to jump over it by simultaneously lifting both feet from the floor and propelling his/her body forward and landing with feet together.)

Acceptable Response

-- Successful performance of the above activity.

(Mark scoring sheet accordingly.)

53. (Given a directive (opportunity) to skip, as a participant in any group activity which involves skipping, the learner will be able to skip using alternate feet, for a distance of ten or more feet.)

Acceptable Response

—Successful performance of the above activity.

(Mark scoring sheet accordingly.)

---

54. (Given a ten-foot length of a 2" by 4" piece of lumber the child will be able to walk a distance of at least five feet on the 4" side of the lumber.)

Acceptable Response

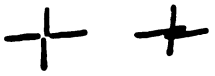
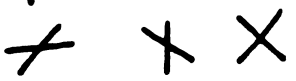
—Successful performance of the above activity.

(Mark your scoring sheet accordingly and put the materials away. Thank and regard the child for working with you.)

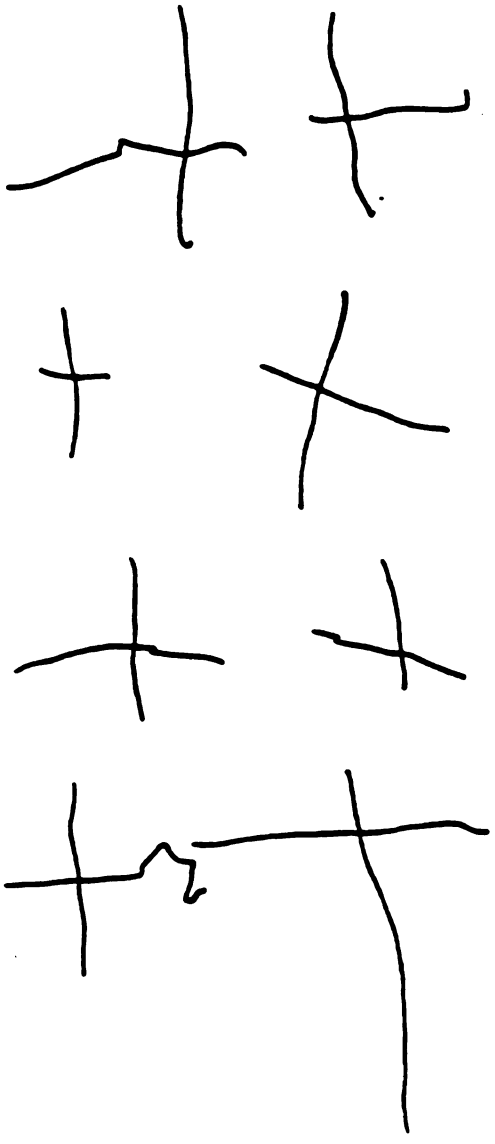
FORM 4 Vertical-Horizontal Cross

Scoring Criteria

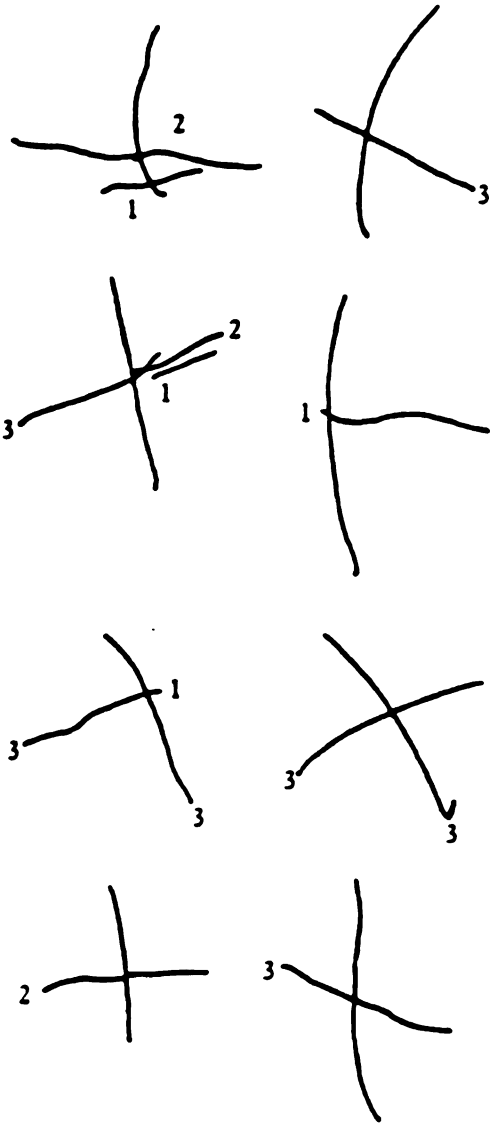
Age Norms
Male: 4-1
Female: 3-8

1. Two fully intersecting lines	not:	
2. Two continuous lines	not:	
3. At least ½ of each line within 20° of its correct orientation	not:	

Passing



Failing



FORM 6 Square

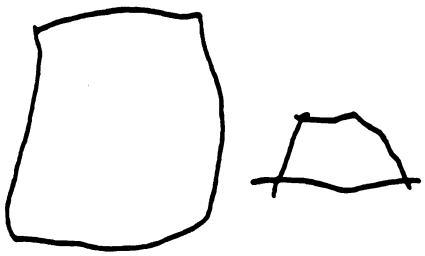
Scoring Criteria

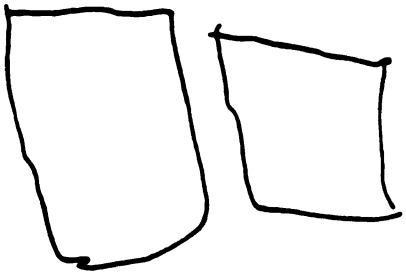
Age Norms
Male: 4-6
Female: 4-3

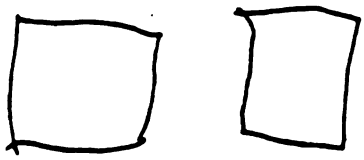
Four clearly defined *sides*
(corners need not be angular)

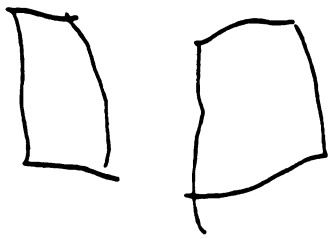
not:

Passing

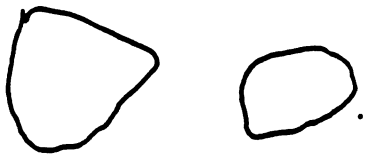


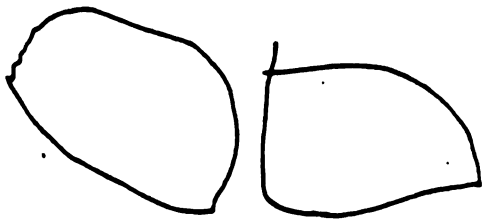


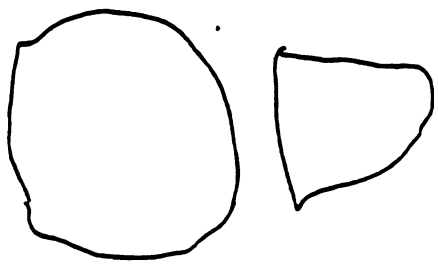




Failing



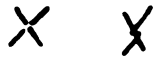




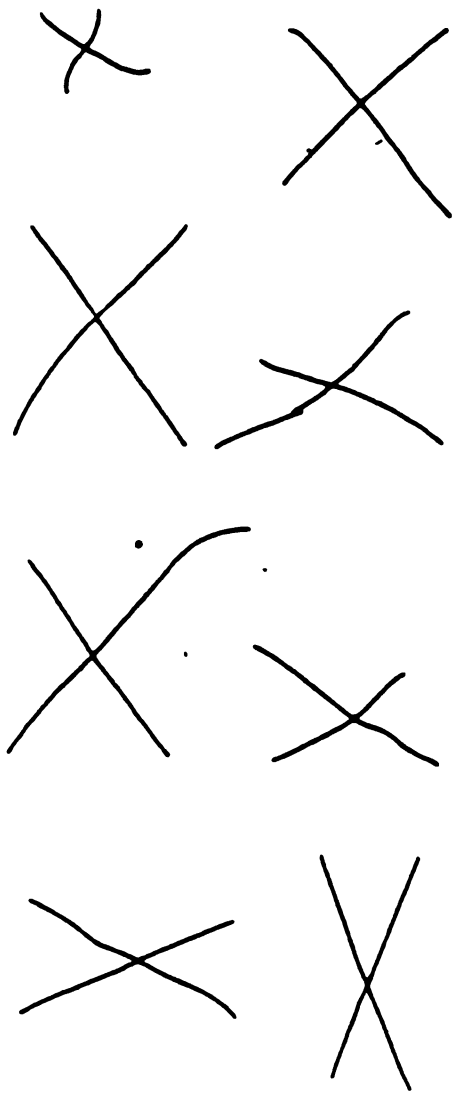
FORM 8 Oblique Cross

Scoring Criteria

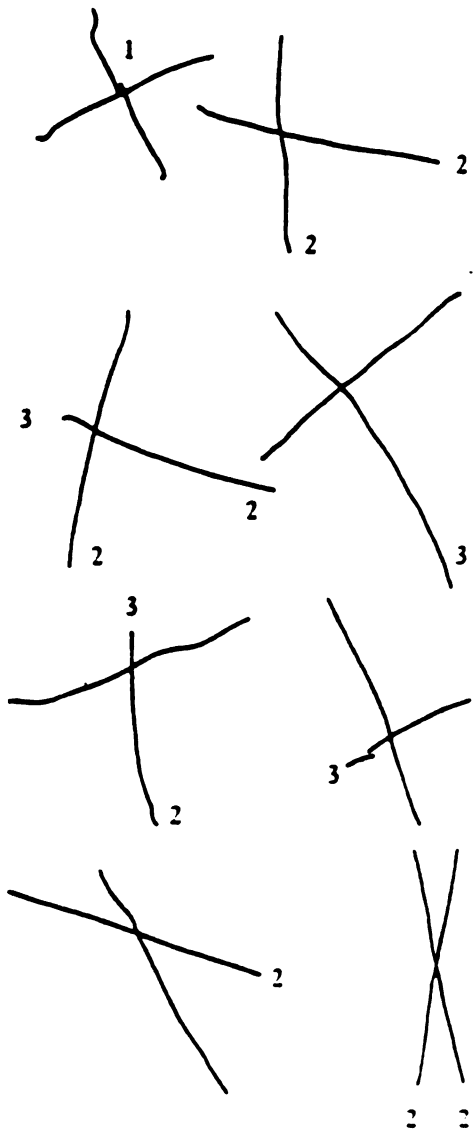
Age Norms
Male: +11
Female: +10

1. Two continuous, intersecting lines — not: 
2. Lines angled between 20°–70° and 110°–160° — not: 
3. Fairly equal length of "leg" — not: 

Passing



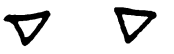
Failing



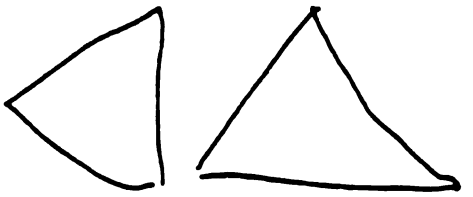
FORM 9 Triangle

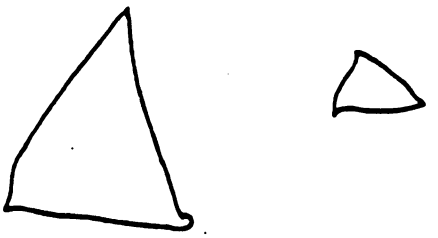
Scoring Criteria

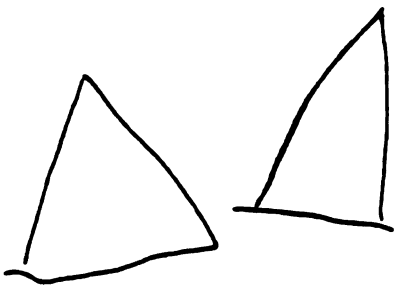
Age Norms
Male: 5-3
Female: 5-3

1. Three clearly defined sides — not: 
2. One corner higher than others — not: 

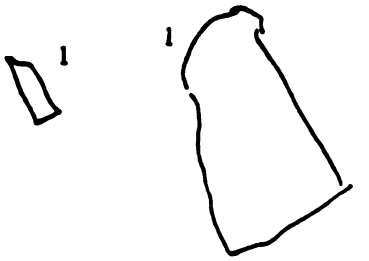
Passing

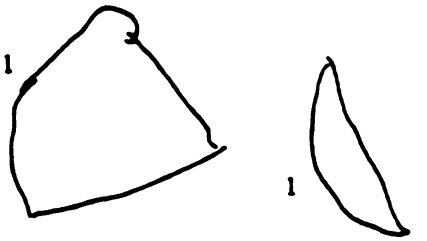


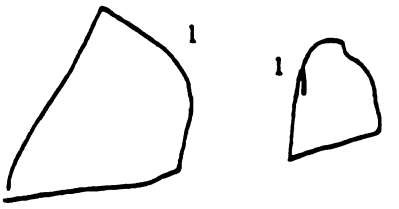




Failing







APPENDIX I

MICHIGAN STATE UNIVERSITY

UNIVERSITY COMMITTEE ON RESEARCH INVOLVING
HUMAN SUBJECTS (UCRIHS)
238 ADMINISTRATION BUILDING
(517) 355-2186

EAST LANSING • MICHIGAN • 48824

February 12, 1981

Mr. Richard N. Claus
3207 Curtis Road
Birch Run, Michigan 48415

Dear Mr. Claus:

Subject: Proposal Entitled, "The Development and Evaluation of Prekindergarten Rediness Screening Device"

The above referenced project was recently submitted for review to the UCRIHS.

We are pleased to advise that the rights and welfare of the human subjects appear to be adequately protected and the Committee, therefore, approved this project at its meeting on February 2, 1981.

Projects involving the use of human subjects must be reviewed at least annually. If you plan to continue this project beyond one year, please make provisions for obtaining appropriate UCRIHS approval prior to the anniversary date noted above.

Thank you for bringing this project to our attention. If we can be of any future help, please do not hesitate to let us know.

Sincerely,

Henry E. Bredeck
Chairman, UCRIHS

HEB/jms

cc: Dr. Robert L. Ebel

SCHOOL DISTRICT
OF THE CITY OF
SAGINAW
550 Millard Street
Saginaw, Michigan 48607
517-776-0200
FOSTER B. GIBBS, Ph. D., Superintendent

January 19, 1981

Professor Henry E. Bredeck, Chairman
University Committee on Research
Involving Human Subjects (UCRIHS)
Michigan State University
238 Administration Building
East Lansing, MI 48824

Dear Professor Bredeck:

I have reviewed in detail the research proposal written by Richard N. Claus. The Policy Manual, written by the Board of Education of the School District of the City of Saginaw, stipulates the steps needed to be taken to gain approval to conduct research within this district (see attachment for a copy of the Research Policy). As the key administrator in implementing the policy, I can assure you that Mr. Claus's proposal adequately meets all of our standards to protect human subjects.

If you have any questions, please call.

Sincerely,

ADMINISTRATION AND PERSONNEL

Barry E. Quimper

Barry E. Quimper, Director
Evaluation, Testing and Research

BEQ/kjm

Attachment

RESEARCH POLICY

Any internal or external research involving students, teachers, or other employees of the School District of the City of Saginaw, must follow the procedures outlined below.

The researcher will supply the Director of Testing and Research with the information described below, follow the sequence outlined below, and obtain written permission as stated prior to conducting the research.

1. Describe the implications and benefits of the study for local education.

2. Submit a written plan of action defining the following:

    a. Purpose of the study.
    b. Population involved.
    c. Design of the study, including timeline.
    d. Instrumentation.
    e. Guarantees of confidentiality and anonymity.
    f. Written permission of participants, when appropriate.
    g. Guarantee of feedback to participants.

3. Receive written permission from the Director of Testing and Research to conduct the research.

*Taken from A Policy Manual for the Operation of the Saginaw Evaluation, Testing and Research Department.

APPENDIX J

TABLE J.1. AGE AND SEX OF CHILDREN SCREENED USING THE PREKINDERGARTN READINESS SCREENING DEVICE (PRSD) FOR THREE CONSECUTIVE YEARS--SEPTEMBER/OCTOBER 1978, 1979, AND 1980.

AGE		NUMBER OF CHILDREN SCREENED															
		1978				1979				1980				Three Consecutive Year Total			
Year	Month	Male	Female	Other*	Total	Male	Female	Other	Total	Male	Female	Other	Total	Male	Female	Other	Total
2	1	--	--	--	--	--	1	--	1	--	--	--	--	--	1	--	1
3	0	--	1	--	1	--	--	--	--	--	--	--	--	--	1	--	1
3	6	--	--	--	--	--	1	--	1	--	--	--	--	--	1	--	1
3	8	1	--	--	1	--	--	--	--	--	--	--	--	1	--	--	1
3	9	4	4	--	8	3	4	--	7	--	3	--	3	7	11	--	18
3	10	21	16	--	37	21	14	--	35	12	16	--	28	54	46	--	100
3	11	14	15	1	30	18	23	1	42	14	16	--	30	46	54	2	102
4	0	27	24	--	51	20	19	2	41	15	23	1	39	62	66	3	131
4	1	19	25	1	45	11	18	2	31	21	20	--	41	51	63	3	117
4	2	21	15	2	38	19	28	1	48	26	20	2	48	66	63	5	134
4	3	16	27	--	43	17	24	6	47	16	14	--	30	49	65	6	120
4	4	23	19	--	42	15	15	5	35	13	11	--	24	51	45	5	101
4	5	20	21	--	41	19	22	2	43	21	21	1	43	60	64	3	127
4	6	22	22	--	44	12	26	2	40	16	15	1	32	50	63	3	116
4	7	10	19	--	29	14	15	1	30	19	14	--	33	43	48	1	92
4	8	20	24	--	44	26	18	6	50	12	14	--	26	58	56	6	120
4	9	18	22	--	40	18	11	3	32	12	12	--	24	48	45	3	96
4	10	--	1	--	1	2	1	--	3	2	1	--	3	4	3	--	7
4	11	--	--	--	--	--	2	--	2	1	1	--	2	1	3	--	4
5	0	1	--	--	1	1	--	1	2	--	--	1	1	2	--	2	4
5	2	--	--	--	--	--	2	--	2	--	--	--	--	--	2	--	2
5	3	--	--	--	--	1	--	1	2	--	--	--	--	1	--	1	2
5	5	--	--	--	--	--	--	1	1	--	--	--	--	--	--	1	1
5	6	--	--	--	--	1	--	--	1	1	--	--	1	2	--	--	2
5	8	2	1	--	3	--	--	--	--	--	--	--	--	2	1	--	3
7	1	--	--	--	--	--	--	--	--	--	1	--	1	--	1	--	1
Other		--	1	--	1	1	6	--	7	1	2	--	3	2	9	--	11
TOTAL		239	257	4	500	219	250	34	503	202	204	6	412	660	711	44	1,415

*Incomplete information made it impossible to assign these individuals to another cell.

APPENDIX K

TABLE K.1. RACIAL ETHNIC BACKGROUND OF CHILDREN SCREENED USING THE PREKINDERGARTEN READINESS SCREENING DEVICE (PRSD) FOR THREE CONSECUTIVE YEARS--SEPTEMBER/OCTOBER OF 1978, 1979, AND 1980.

Racial Ethnic Category	Children Screened							
	1978		1979		1980		Total	
	N	%	N	%	N	%	N	%
American Indian or Alaskan Native or Native American	3	0.6	0	0.0	0	0.0	3	0.2
White, Not of Latino or Hispanic Origin	27	5.4	38	7.6	23	5.6	88	6.2
Latino or Hispanic	78	15.6	84	16.7	69	16.7	231	16.3
Black, Not of Latino or Hispanic Origin	386	77.2	378	75.1	319	77.4	1083	76.5
Asian or Pacific Islanders	0	0.0	1	0.2	1	0.2	2	0.1
Other*	6	1.2	2	0.4	0	0.0	8	0.6
Total	500	100.0	503	100.0	412	99.9**	1415	99.9**

*No race data was provided for these individuals.
**Due to rounding error.

APPENDIX L

TABLE L.1. DEMOGRAPHIC DATA FROM THE 1970 CENSUS SHOWING CHARACTERISTICS OF FAMILIES AND HOMES IN THE TITLE I ATTENDANCE AREA OF THE SCHOOL DISTRICT OF THE CITY OF SAGINAW, MICHIGAN.

Characteristic	Range
Range of Family Income	$4,000 to $10,400
Percent of Families Below Poverty Level	11% to 45%
Percent of Parents Having a High School Education	18% to 40%
Percent of Parents Having a College Education	1.1% to 6.7%
Percent of Children from Homes in a Husband-Wife Household	38% to 82%
Percent of Children from Homes Where the Head of the Household is Female	14% to 50%

APPENDIX M

FIGURE M.1.
SPEARMAN-BROWN SPLIT HALF FORMULA FOR ESTIMATING TOTAL TEST RELIABILITY FROM PARTS OF UNEQUAL LENGTH

$$R = \frac{r\sqrt{r^2 + 4\,Pq\,(1-r^2)} - r}{2\,Pq\,(1-r^2)}$$

Legend

R = Estimated reliability of total test
P = Proportion of total testing time taken up by part P
q = Proportion of total testing time taken up by part q
r = Correlation between parts P and q

APPENDIX N

TABLE N.1. NUMBER AND PERCENT OF AN AGE GROUP PASSING EACH PRSD TEST ITEM.

PRSD Item Number	Behavior Requested	Age Group (Years-Months) 4-0 or less N=355	4-1 to 4-3 N=371	4-4 to 4-6 N=344	4-7 or more N=334
1	Say age	N 166 % 46.76	209 56.33	197 57.26	187 55.98
2	Say name	N 202 % 56.90	216 58.22	234 68.02	251 75.14
3	Point to neck	N 245 % 69.01	281 75.74	265 77.03	287 85.92
4	Identify body part	N 87 % 24.05	104 28.03	139 40.40	137 41.01
5	Tell function of bodypart	N 40 % 11.26	58 15.63	91 26.45	89 26.64
6	Pick up same color	N 220 % 61.97	266 71.69	272 79.06	268 80.23
7	Pick up named color	N 168 % 47.32	191 51.48	212 61.62	218 65.26
8	Say color of object	N 118 % 33.39	137 36.92	163 47.38	182 54.49
9	Say color of object	N 153 % 43.09	163 43.93	202 58.72	208 62.27
10	Count to five	N 111 % 31.26	126 33.96	144 41.86	158 47.30
11	Give me four blocks	N 39 % 10.98	41 11.05	69 20.05	83 24.85
12	Say number of blocks	N 112 % 31.54	143 38.54	159 46.22	167 50.00
13	Pick up blocks	N 278 % 78.30	293 78.97	290 84.30	282 84.43
14	Walk backwards	N 72 % 20.28	84 22.64	91 26.45	107 32.03

TABLE N.1 (CONTINUED)

PRSD Item Number	Behavior Requested	Age Group (Years-Months) 4-0 or less N=355	4-1 to 4-3 N=371	4-4 to 4-6 N=344	4-7 or more N=334
15	Carry out 3-part command	N 44	51	74	77
		% 12.39	13.74	21.51	23.05
16	Tell what books are for	N 163	195	227	229
		% 45.91	52.56	65.98	68.56
17	Draw cross given model	N 53	96	102	126
		% 14.92	25.87	29.65	37.72
18	Hop after demonstrated	N 183	211	229	229
		% 51.54	56.87	66.56	68.56
19	Throw ball five feet	N 261	282	261	270
		% 73.52	76.01	75.87	80.83
20	Point to shape like a wheel	N 244	261	249	260
		% 68.73	70.35	72.38	77.84
21	Point to shape like a tent	N 168	206	197	212
		% 47.32	55.52	57.26	63.47
22	Point to shape like a stick	N 199	240	223	242
		% 56.05	64.69	64.82	72.45
23	Tell which is bigger	N 192	228	219	229
		% 54.08	61.45	63.66	68.56
24	Draw diamond given model	N 43	70	99	119
		% 12.11	18.86	28.77	35.62
25	Tell nursery rhyme, song, poem	N 53	54	79	75
		% 14.92	14.55	22.96	22.45
26	Say yellow jello	N 167	207	192	220
		% 47.04	55.79	55.81	65.86
27	More than one word response	N 99	87	92	104
		% 27.88	23.45	26.74	31.13

LIST OF REFERENCES

LIST OF REFERENCES

Almy, M. Young Children's Thinking. New York: Teacher's College Press, Teacher's College, Columbia University, 1966.

Anastasi, A. Standardized Ability Testing. In P.H. Mussen (Ed.), Handbook of Research Methods in Child Development. New York: Wiley, 1970.

Anderson, R., Miles, M., and Matheny, P. Communicative Evaluation Chart. Cambridge, MA: Educators Publishing Services, Inc., 1964.

Banham, K.M. Quick Screening Scale of Mental Development. Munster, IN: Psychological Affiliates, 1963.

Bayley, N. Bayley Scales of Infant Development. New York: The Psychological Corp., 1969.

Beery, K.E. and Buktenica, N.A. Developmental Test Visual-Motor Integration. Chicago, IL: Follett Publishing Co., 1967.

Bennett, E.L., Diamond, M.C., Kretch, D. and Rosenzweig, M.R. Chemical and Anatomical Plasticity of the Brain. Science, 1964, 146, 610-619.

Bloom, B.S. Stability and Change in Human Characteristics. New York: Wiley, 1964.

Borg, W.R. and Gall, M.D. Educational Research: An Introduction (2nd ed.). New York: David McKay Company, Inc., 1971.

Bruner, J.S. The Process of Education. New York: Vintage Books, 1960.

Buros, O.K. (Ed.). The Sixth Mental Measurement Yearbook. Highland Park, N.J.: Gryphon Press, 1965.

Buros, O.K. (Ed.). The Seventh Mental Measurement Yearbook (Vol. 1 & 2). Highland Park, N.J.: Gryphon Press, 1972.

Buros, O.K. (Ed.). The Eighth Mental Measurement Yearbook (Vol. 1 & 2). Highland Park, N.J.: Gryphon Press, 1978.

Caldwell, B.M. The Preschool Inventory. Princeton, N.J.: Educational Testing Service, 1967.

Cattell, P. Measurement of Intelligence of Infants and Young Children (rev. ed.). New York: The Psychological Corp., 1960.

Cazden, C.B. Evaluation of Learning In Preschool Education: Early Language Development. In B.S. Bloom, J.T. Hastings, and G.F. Madaus (Eds.), Handbook on Formative and Summative Evaluation of Student Learning. New York: McGraw Hill Book Company, 1971.

Chapman, J.E. and Lazar, J.B. The Present Status and Future Needs in Day Care Research. Washington, D.C.: Education Resources Division Capitol Publications, Inc., 1971.

Danger, V.A., Lyon, T.M., and Gerber, M.F. Daberon: A Screening Device for School Readiness. Portland, OR: Daberon Research; 1972.

Doll, E.A. Preschool Attainment Record. Circle Pines, MN: American Guidance Service, 1965.

Doll, E.A. Vineland Social Maturity Scale. Circle Pines, MN: American Guidance Service, 1965.

Dowley, E.M. Early Childhood Education. In R.L. Ebel and V.H. Noll (Eds.), Encyclopedia of Educational Research (4th ed.). New York: MacMillan Company, 1969.

Ebel, R.L. Estimation of the Reliability of Ratings. Psychometrika, 1951, 16 (4), 407-424.

Fallon, B.J. (Ed.). 40 Innovative Programs in Early Childhood Education. Belmont, CA: Lear Siegler, Inc./ Fearon Publishers, 1973.

Frank, L.K. Fundamental Needs of the Child. Mental Hygiene, 1938, 22, 353-379.

Frankenburg, W.K. and Dodds, J.B. Denver Developmental Screening Test. Denver, CO: University of Colorado Medical Center, 1969.

Gagne, R.M. The Conditions of Learning (2nd ed.). New York: Holt, Rinehart, and Winston, 1970.

Gesell, A., et al. The First Five Years of Life: A Guide to the Study of the Preschool Child. New York: Harper, 1940.

Gesell, A., et al. Infant and Child in the Culture of Today: The Guidance of Development in Home and Nursery School. New York: Harper, 1943.

Gesell Institute of Human Development. Preschool Test. Lumberville, PA, 1979.

Goodenough, F.L., Maurer, K.M., and Van Wagenen, M.J. Minnesota Preschool Scale. Minneapolis: Educational Test Bureau, 1940.

Goodland, J.I., Klein, M.F., and Novotney, J.M. Early Schooling in the United States. New York: McGraw-Hill Book Company, 1973.

Goodwin, W.L. and Driscoll, L.A. Handbook of Measurement and Evaluation in Early Childhood Education. San Francisco: Jossey-Bass Publishers, 1980.

Grotberg, E.H. Review of Research 1965 to 1969 (OEO Pamphlet, 6108-13). Research and Evaluation Office Project Head Start, Office of Economic Opportunity, June, 1969.

Guilford, J.P. Psychometric Methods (2nd ed.). New York: McGraw-Hill Book Company, 1954.

Guilford, J.P. and Fruchter, B. Fundamental Statistics in Psychology and Education (5th ed.). New York: McGraw-Hill Book Company, 1973.

Hall, E.T. The Hidden Dimension. New York: Doubleday and Company, 1966.

Hall, E.T. The Silent Language. New York: Doubleday and Company, 1965.

Hess, R.D. and Bear, R.M. (Eds.). Early Education. Chicago: Aldine, 1968.

Hoepfner, R., Stern, C., and Nummedal, S.G. (Eds.). CSE-ECRC Preschool Kindergarten Test Evaluations. Los Angeles: UCLA Graduate School of Education, 1971.

Horrocks, J.E. and Schoonover, T.I. Measurement for Teachers. Columbus, OH: Merrill, 1968.

Horst, P. Estimating Total Test Reliability from Parts of Unequal Length. Educational and Psychological Measurement, 1951, 11 (3), 368-371.

Hunt, J.McV. Intelligence and Experience. New York: Ronald Press, 1961.

Isaac, S. Handbook in Research and Evaluation. San Diego, CA: Knapp, 1971.

Ilg, F.L. and Ames, L.B. School Readiness: Behavior Tests Used at the Gessell Institute. New York: Harper and Row, Publishers, 1965.

Johnson, O.G. and Bommarito, J.W. Tests and Measurements in Child Development: A Handbook. San Francisco: Jossey-Bass Publishers, 1971.

Kamii, C.K. Evaluation of Learning in Preschool Education: Socio-emotional, Perceptual-motor, and Cognitive Development. In B.S. Bloom, J.T. Hastings, and G.F. Madaus (Eds.), Handbook on Formative and Summative Evaluation of Student Learning. New York: McGraw-Hill Book Company, 1971.

Kirk, S.A. Early Education of the Mentally Retarded. Urbana: University of Illinois Press, 1958.

Lazerson, M. Social Reform and Early Childhood Education: Some Historical Perspectives. In R.H. Anderson and H.G. Shane (Eds.), As the Twig is Bent: Readings in Early Education. Boston: Houghton Mifflin Company, 1971.

Lazar, I. and Darlington, R. Lasting Effects After Preschool. (Summary Report, DHEW Publication No./OHDS/ 79-30179). Washington, D.C.: U.S. Government Printing Office, 1979.

Lesiak, W.J., Jr. Developmental Tasks for Kindergarten Readiness. (Archives of Behavioral Science Monograph No. 52). Brandon, VT: Clinical Psychology Publishing Company, Inc., 1978.

Maitland, S., Nadeau, J.B.E., and Nadeau, G. Early School Screening Practices. Journal of Learning Disabilities. 1974, 7 (10), 55-58.

Marascuilo, L.A. and McSweeney, M. Nonparametric and Distribution-Free Methods for the Social Sciences. Monterey, CA: Brooks/Cole Publishing Company, 1977.

Mehrens, W.A. and Lehmann, I.J. Measurement and Evaluation in Education and Psychology. New York: Holt, Rinehart and Winston, Inc., 1973.

National School Public Relations Association (Ed.). Preschool Breakthrough: What Works in Early Childhood Education. (Education USA Special Report) Arlington, VA: Editor, 1970.

National School Public Relations Association (Ed.). Early Childhood Education: Current Trends in School Policies and Programs. (Education USA Special Report) Arlington, VA: Editor, 1973.

Osborn, D.K. Early Child Education in Historical Perspective. Athens, GA: Education Associates, 1975.

Preschool Education Pays Off For Poor Students, Study Says. Education USA, December 22, 1980, pp. 129; 132.

Preschool Programs Cost Effective, Riles Says. Report on Education Research, December 10, 1980, pp. 9-10.

Preschool Research Shows Lasting Benefit. Report on Education Research, December 24, 1980, p. 2.

Prescott, G.A., Balow, I.H., Hogan, T.P., and Farr, R.C. Metropolitan Achievement Tests: Teacher's Manual for Administering and Interpreting. New York: Harcourt Brace Jovanovich, Inc., 1978.

Quick, A.D., Little, T.L., and Campbell, A.A. MEMPHIS Comprehensive Developmental Scale. Belmont, CA: Fearon Publishers, Inc., 1974.

Remmers, H.H. Rating Methods in Research on Teaching. In N.L. Gage (Ed.), Handbook of Research on Teaching. Chicago: Rand McNally, 1963.

Sears, P.S. and Dowley, E.M. Research on Teaching in the Nursery School. In N.L. Gage (Ed.), Handbook on Research in Teaching. Chicago: Rand McNally, 1963.

Shane, H.G. The Renaissance of Early Childhood Education. In R.H. Anderson and H.G. Shane (Eds.), As the Twig is Bent: Readings in Early Childhood Education. Boston: Houghton Mifflin Company, 1971.

Sheridan, M.D. The Developmental Progress of Infants and Young Children. London: H.M.S.O., 1968.

Slater, S.J. Applied Performance Testing: Typology, Advances, Limitations, and Examples. Unpublished paper presented at the Annual Meeting of the American Educational Research Association, Boston, 1980.

Slosson, L. Slosson Intelligence Test. East Aurora, NY: Slosson Education Publications, 1964.

Spock, B. Common Sense Book of Baby and Child Care. New York: Duell, Sloan, and Pearce, 1946.

Stanley, J.C. (Ed.). Preschool Programs for the Disadvantaged: Five Experimental Approaches to Early Childhood Education. Baltimore, MD: The John Hopkins University Press, 1972.

Stolz, L.M., et al. Father Relations of War-Born Children. Stanford, CA: Stanford University Press, 1954.

Stutsman, R. Merrill-Palmer Scale. Detroit: Merrill-Palmer School, 1963.

The Ups and the Downs of Preschool Intervention. Education USA, May 12, 1980, p. 285.

Walker, W. A Readiness Test for Disadvantaged Preschool Children. (DHEW Publication No. /OE/72-9) Washington, D.C.: U.S. Government Printing Office, 1972.

Whitely, S. and Davis, R.V. The Nature of Objectivity With the Rasch Model. Journal of Educational Measurement, 1974, 11, 163-178.

Winer, B.J. Statistical Principles in Experimental Design (2nd ed.). New York: McGraw-Hill Book Company, 1971.

Wolfe, D.A. and Hollander, M. Nonparametric Statistical Methods. New York: Wiley, 1973.

Zehrback, R.R. Comprehensive Identification Process, University of Illinois at Urbana-Champaign: Institute for Child Behavior and Development, 1975.

Zeitlin, S. Kindergarten Screening: Early Identification of Potential High Risk Learners. Springfield, IL: Thomas, 1976.

Zimmerman, I., Steiner, V., and Evatt, R. Preschool Language Scale. Columbus, OH: Merrill, 1969.