5-1975

A Comparison of Second Grade Children's Learning Curves on School Tasks with Their Respective Performances on the "Black Box Test of Learning Ability"

Robert Stephen Knox
Utah State University

A COMPARISON OF SECOND GRADE CHILDREN'S LEARNING CURVES ON SCHOOL TASKS WITH THEIR RESPECTIVE PERFORMANCES ON THE "BLACK BOX TEST OF LEARNING ABILITY"

by

Robert Stephen Knox

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Psychology

Approved:

UTAH STATE UNIVERSITY
Logan, Utah

1975

ACKNOWLEDGMENTS

I would like to take this opportunity to thank the following members of my committee for their helpful suggestions and general support: Dr. Wayne Wright, Dr. Devoe Rickert, Dr. Marvin Fifield, and Dr. Michael Bertoch. The support and continued assistance throughout the entire process, provided by Dr. Glendon Casto, chairman of this dissertation, is sincerely appreciated.

I entend my sincere thanks to Dr. Evelyn Thirkill for granting me permission and the time to work on this project, and to Mrs. Nicky Swisher, Mrs. Patsy Dunn, and Mr. Roger Wheeler, personnel of Tendoy Elementary School, whose assistance and permission made this study possible.

Finally, the author would like to thank Sheila Ann Knox for her valuable assistance in the final preparation of the figures.

Robert Stephen Knox

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

A Comparison of Second Grade Children's Learning Curves on School Tasks with Their Respective Performances on the "Black Box Test of Learning Ability"

by

Robert Stephen Knox

Utah State University, 1975

Major Professor: E. Wayne Wright, Ed.D.
Dissertation Chairman: Glendon Casto, Ph.D.
Department: Psychology

This study investigated the usefulness of the Black Box Test of Learning Ability as an indicator of children's performance on math and writing tasks. Twelve second grade students, seven to eight years of age, demonstrated naivete on both tasks and were subsequently individually administered the learning test. The subjects were divided into two groups, and each group received a different task presentation order. Composite scores were derived for all subjects and tasks, and individual learning curves were compared.

The combined Black Test score produced a significantly high correlate to math (rho = .733) and writing (rho = .841) than either the paired associate or sequential learning tasks alone. Although the length of the learning curve of the BBT was indicative of the learning curve lengths for each school task, trial by trial learning curve comparisons were less reliable. Analysis of these data suggested that the method of instruction (i.e., the amount of attempted practice and appropriate feedback) was the major factor correlating the Black Box Test to each school task. It was suggested that the most useful assessment of "learning ability" would employ a behavioral sample of the task itself, rather than utilize a correlated activity.

(134 pages)

INTRODUCTION

Within the last 70 years, the field of education has dedicated a major portion of its research and development efforts towards estimating children's learning ability within the school environment. Historically, this orientation has led to the development of intelligence tests as a primary mode of assessing an individual's potential for learning. Unfortunately, the results of such tests are often directly related to the concept of "intelligence" (viewed as a separate entity consisting of mental processes which control the potential for future learning) rather than merely as a sample of one's intelligent behavior (measuring the retention of specific knowledge and skills previously learned). Presently, the construct of intelligence has neither been adequately nor consistently defined and therefore remains a source of theoretical disagreement among its major proponents and of limited usefulness in the educational setting.

The most widely acclaimed initiators of the intelligence tests were Binet and Simon, who, in 1905, created a screening device consisting of 30 problems of increasing difficulty. Their purpose was to assist the Paris educational system in selecting children which would profit from academic instruction. It is interesting to note that, although this device was created and utilized nearly 60 years ago for such a specific purpose, few basic changes are apparent within intelligence tests currently being used (Mussen, Conger, and Kagen, 1969). Since 1905 to the present, the major use of intelligence tests has resulted in a system for predicting future grades, rather than as a tool for providing information concerning an individual's learning style and subsequently as an accurate indicator of true learning ability.

The shortcomings of intelligence tests as devices for assessing learning ability are further supported by the lack of continuity between the definition of intelligence and the devices utilized to measure it. Although some discrepancies exist in defining intelligence, many definitions either imply or clearly state that it is best understood as the assessment of an individual's ability to learn (Goodenough, 1959; Hoveland, 1959; McGeoch, 1942; Drever, 1965; and Mussen, Conger, and Kagan, 1969). Leaders in the field such as Piaget and Guilford have suggested that intelligence is best defined as the "ease with which a child learns a new idea or a new set of behaviors." Devices commonly used to measure intelligence, such as the Weschler Intelligence Scale or the Stanford-Binet, however, infrequently assess a child's ability to learn a new skill or fact. Rather, they measure a skill or a piece of knowledge the child has already learned. Developers of intelligence tests have apparently operated under the premise that the best method for assessing learning ability is to measure the individual's amount of previously learned knowledge. The result of such an assumption may be of little educational value since neither the reasons for an individual's lack of specific knowledge nor information concerning the training necessary to produce learning is provided.

Another area of difficulty often encountered with previously established intelligence tests as measurements of learning ability is that they consist of unequal items on the familiarity-unfamiliarity continuum. Havighurst (1951) stated:

> . . . to be equally fair to all persons, an intelligence test should present problems which are either equally familiar or equally unfamiliar [p. 16].

Numerous studies are cited throughout the literature which demonstrate that intelligence tests are highly susceptible to differences between cultures and environments. Many of these studies demonstrated that children from an urban-type background produced higher scores than children of comparable age from the rural setting (e.g., Gaw, 1925; Fahny, 1954; and Lehmann, 1959). However, studies such as Gaw (1925) and Fahny (1954) additionally demonstrated that although children from rural settings produced lower overall intelligence test scores when compared with children from urban settings, their performances surpassed that of urban children on those items similar to their own cultural experience. Such studies indicate that intelligence tests primarily measure the differences between individuals' amount of exposure to numerous stimuli.

Other studies have demonstrated that the score derived from an intelligence test is often related to environmental changes. Klineberg (1935) and Lee (1951), while investigating the migration of black children from rural to urban areas, demonstrated that IQ scores increase as the length of residence in urban areas increases. Sherman and Key (1932), who studied isolated mountain children, suggest that the inverse is also true. Their results showed that IQ decreases as the length of residence in a rural area increases. Rosandic and Bakvic (1970) demonstrated that equalizing the conditions of intellectual development (schools) reduces significantly the differences in IQ scores between children originating from favorable as opposed to unfavorable environments. Therefore, a question is raised as to the ability of current intelligence tests to measure individual learning potential. It would seem that the measurement of such a potential should be less influenced by cultural changes and experience.

Although intelligence tests have frequently been shown to produce relatively high correlations between IQ scores and overall school achievement (Stake, 1961), other studies (e.g., Mitchell, 1943) reported correlations as low as .15. Bond (1940) demonstrated that correlation factors fluctuate radically between IQ scores and specific areas of school achievement. According to Anastasi (1961), the correlations between the Stanford-Binet IQ and school achievement (measured by grades, teacher ratings, and achievement tests) fluctuate between .40 and .75, depending upon the particular school subject. Clelland and Toussaint (1962) reported correlations as low as .21 between IQ scores and arithmetic.

Besides the inconsistencies which exist between studies attempting to demonstrate the usefulness of intelligence tests as predictors of school achievement, Becker, Englemann and Thomas (1971) raised the following considerations:

> . . . a major criterion in the selection of items for Binet-type IQ tests is the correlations of items with progress in school ('learning faster'). Thus, it would be reasonable to call the IQ test a test of school progress. It is then possible to conclude that school progress effects school progress. The statement is, of course, tautological. . . .
>
> . . . If the IQ test measures concepts and operations that the teacher assumes the children already know, it is implied that the children who came to school knowing less are taught less. If the teacher teaches to the average student, she is bound to lose (not teach) the children who are not prepared to follow her lesson. The children who already know what the teacher is teaching will appear to be smart and 'catch on quickly' [p. 393].

In summary, it is concluded that the intelligence test is far from an adequate device for measuring an individual's learning ability, particularly as it relates to measuring a child's potential for learning somewhat novel tasks in the school environment. One of the major reasons for the administration of intelligence tests in education, in spite of its continual criticism and frequent misuse, is that it readily lends itself to a system of classification. This is particularly advantageous for the administration of school since state appropriations are typically awarded to local school districts on this basis. Although the classification of students will probably exist for many years, a new emphasis is being demanded by both state and governmental agencies, not only to educate but to provide quality education for <u>all</u> children including those considered "exceptional." According to Schrag (1973)

> . . . the United States Commissioner of Education called for a national goal of full educational opportunities for all exceptional children by 1980 [p. 1].

Rationale for the Present Study

It has been proposed that educators need to concern themselves more with the process of learning than with merely predicting one's potential for future academic performance. Thus, it is felt that more information relating to the learning process would be provided by a test of learning ability which produces a learning curve by measuring ongoing behavior change, than is now obtained by the use of intelligence tests, which merely provide total scores by measuring previously learned behavior. Specific information regarding a child's learning style (e.g., lengths of the curve, trial by trial performance levels) could assist in making instructional decisions relating to grouping, the need for frequent "appropriate practice" and "immediate feedback," types of materials, programming formats, etc.

Although the need for such information is great, most ability tests presently utilized in education do not provide data which readily assist in making various instructional recommendations. Demonstrating that children learn novel tasks, but at different speeds, rather than emphasizing what they have or have not learned (whether taught or not) may also facilitate positive changes in teachers' attitude.

Although previous research efforts have emphasized the value of investigating the relationship between learning tasks and academic achievement (i.e., Wallen, 1959; Stakes, 1961; Meuller, 1968; and Green and Rohwer, 1971), these studies typically utilized tasks which merely assessed the amount of learning in a predetermined number of study-test sessions. The major criticism of this approach is that valuable learning data are lost, since the children are not required to learn the entire task and information concerning the final learning performance is never tapped. It would seem that, in order to provide the maximum information concerning a student's learning ability, the learning task should: 1) provide a completely novel set of tasks, 2) contain stimuli and problems which are universally unfamiliar to all subjects, 3) provide information concerning how fast a child can learn something new within the test situation, 4) provide data which can be readily transformed into composite scores as well as learning curves, and 5) provide a format specifying task completion contingent upon an accuracy criterion, rather than the number of study-test sessions administered.

In 1972, Sam Campanella developed a device entitled the "Black Box Test of Learning Ability" which meets all of the above qualifications. This test consists of a paired associate and a sequential learning task, designed to measure ongoing behavior change by requiring the subjects to learn a completely novel set of tasks. The stimuli and problems are universally unfamiliar to all subjects, thus providing a possible basis for a culturally-free assessment of learning. Performance data are readily reproduced as learning curves since all subjects are required to participate until the learning criterion is met or a ceiling level is reached (after which the probability of increasing the level of accuracy for that subject is approximately zero). Although some research (Campanella, 1972; and Meeks, 1973) has supported the Black Box Test as a culturally-free assessment device of learning, no research has been undertaken to establish its standardization nor to investigate its usefulness in the educational setting.

Previous research investigating the relationship between learning tasks and academic achievement (e.g., Wallen, 1959; Stakes, 1961; Meuller, 1968; Campanella, 1972; and Meeks, 1973) employed achievement tests, teacher ratings, and grades as indicators of children's school performance. No studies, however, have attempted to relate a test of learning ability to specific academic tasks within the standard elementary school curriculum. It is also interesting to note that those studies investigating the relationship of student's performance on learning tasks as a predictor of achievement, typically omit an investigation of the individual learning curves. Neither have these studies employed a design which allowed for learning curve comparisons between the learning task and the academic skill areas. The reasons for this void throughout the literature appear to be two-fold: 1) the infrequency of studies designed to teach the learning task as well as the specific academic skill to an accuracy criterion within the test situation, and 2) the difficulty involved in collecting comparative data across various tasks.

Finally, few studies have compared the paired associate and sequential learning tasks as predictors of school achievement. Green and Rohwer (1970) suggested that these two learning tasks tap different learning abilities since they did not predict academic achievement (scores on achievement batteries) equally. Jensen (1969), however, stated that learning takes place on two levels, associative (Level I) and conceptual and problem solving (Level II). He further emphasized that sequential and paired associate learning (particularly as it relates to trial and error learning with feedback for correct responses) were both examples of Level I learning ability.

Problem Statement

It was the intent of this study to provide additional information concerning the relationship between learning ability, as measured by the Black Box Test, and the academic performance of second grade children. This study did not intend to establish either the reliability or the validity of the Black Box Test as a predictor of children's learning ability. Rather, it was intended to investigate its relationship to closely associated school tasks, in an effort to determine its future potential as a device for: 1) describing various learning style variables (e.g., learning curve lengths, progressive performance levels) via individually presented learning curves, and 2) estimating school task performance via composite score analysis. Specifically, comparisons were made between performance on the Black Box Test of Learning Ability and children's performance on two traditionally taught academic skill areas within the standard elementary school curriculum (multiplication facts and cursive writing). Both composite scores (based on errors to criterion) and learning curves were provided for each subject across all tasks. Statistical analyses were conducted to determine the correlations between each of the learning task scores (sequential, paired associate, and combined) with each academic skill area. A single subject design and the use of a similar data collection system across all tasks allowed for inter- and intrasubject comparisons.

Hypotheses

1. There will be a significant correlation (.05 level) between all three learning tasks' scores (paired associate, sequential, and combined) measured by errors to criterion, on the Black Box Test of Learning Ability and the learning scores (errors to criterion) on two novel school tasks (multiplication facts and cursive reproduction of eight selected letters of the alphabet).

2. The correlations between the combined learning task score of the Black Box Test and the school tasks will

be higher for that task essentially representing associative learning (memorization of multiplication facts) than with a school task involving the combination of associative learning and fine-motor coordination skills (reproducing selected letters of the alphabet).

REVIEW OF THE LITERATURE

A large number of studies exist in the literature which deal with various assessment devices and their relation to children's future behavior. A majority of these studies primarily investigated the relationship of intelligence test scores with academic success. The difficulties inherent within this type of research development have previously been discussed in Chapter 1. The intent of this review of the literature will, therefore, be limited to: 1) the relevant studies in the literature dealing with the relationship between the assessment of learning ability and other areas of children's performance (typically exhibited by intelligence test scores, achievement batteries, and performance tasks), and 2) a review of the research conducted previously on the Black Box Test of Learning Ability.

There are many differing opinions cited throughout the literature in regard to the specific devices which best assess "learning ability." This review, however, will investigate only those studies utilizing assessment devices which "show how fast a child can learn something relatively new and unfamiliar, right in the test situation" (Jensen, 1969).

Learning Ability Test Performance and Intelligent Test Behavior[1]

Previous research has investigated the relationship of various types of learning tasks with intelligence test behavior and performance on other tests frequently used to estimate intelligence test behavior (i.e., Peabody Picture Vocabulary Test and the Children's Progressive Matrices).

Thompson and Witryol (1942) demonstrated that the ability to learn a motor task was highly correlated with intelligence test behavior (Otis Quick-Scoring Mental Ability Test). Forty unselected volunteers from two elementary psychology classes, 32 women and 8 men, were trained on a high finger-relief maze. The subjects were allowed as many trials as necessary to achieve criterion (three successive runs at 100% accuracy) in one sitting. Three measures of motor learning efficiency were made available: number of trials required to reach criterion, total number of errors,

---

[1]Individuals' performance on tests of intelligence and tests of learning ability will be referred to as "intelligence test behavior" and "learning ability test performance" respectively, in an attempt to avoid confusion as to the meaning of "intelligence" and "learning ability."

and total amount of time. The results of this study demonstrated that .731, .739, and .759 predicted correlational coefficients could be expected between intelligence test behavior and trials, errors, and time, respectively.

Many studies investigating the relationship between these variables have utilized the paired associate technique. Reese and Lipsitt (1970) clearly defined the paired associate task as:

> . . . typically consisting of the presentation of a list of n pairs of items, one pair at a time in different random orders, until the subject has reached some criterion, usually two or three consecutive errorless runs through the list [p. 196].

English and Kidder (1969) employed 27 four year-old kindergarten children and administered the Stanford-Binet Intelligence Test and the Peabody Picture Vocabulary Test to each. All subjects were trained to criterion on a paired associate learning task, consisting of pairing six colors with six pictures. Four second intervals during pairings and between trials were maintained throughout the study and a Hunter Cardmaster Model 340 presented the stimuli. The results demonstrated that although no significant correlations existed between the Peabody Picture Vocabulary Test (PPVT) and the number of trials to learn the paired associate task, a highly significant (.01 level) correlation (r = 0.71) was evident between the intelligent test behavior (Stanford-Binet Intelligence Test) and the learning task. Blue (1970) identified 36 retarded children (determined from previous administrations of the Weschler Intelligence Scale for Children) and 72 children of average "intelligence" (scores ranging from 90 to 110 on the Otis Quick Scoring Mental Ability Test). The latter students were divided into two groups, equated with the retarded students on CA and MA. All subjects were then required to learn, via the paired associate method, 10 paired items. Three types of sensory stimulus presentations were manipulated (visual, auditory, and combined). The learning criterion was established at 10 successive correct responses (100% accuracy) within a single test trial. No significant differences were indicated between the types of sensory-stimulus presentations. The results did demonstrate, however, that "average" subjects were superior in speed of learning to retarded subjects of equal CA. In addition, average subjects reached the learning criterion faster than retarded subjects when equated for mental age. Walther (1969) employed 90 subjects in a study which basically replicated English and Kidder (1969). The results provided additional support to the reported findings of the previous study. Carver and Dubois (1967) expanded the types of learning tasks to be correlated with intelligence test behavior to include: coding, tracing numbers, paired associate learning, and learning two programmed instructional manuals. The intelligence test administered to 269 enlisted men of the United States Navy was the General Classification Test. The results of this study indicated that the performance of these subjects on all of the learning tasks, with the exception of tracing numbers, was significantly correlated (.05 level) to the IQ scores.

The above studies have provided evidence to indicate a strong positive relationship between various learning tasks and intelligence test behavior. These correlations may be expected since intelligence tests measure previously learned behavior (Mussen, Conger, and Kagan, 1969), while tests of learning ability measure "the capacity for acquiring, retaining, and producing new information" (Rohwer, 1971).

A small number of studies, however, indicated that IQ scores are not an accurate reflection of learning proficiency. These studies have typically employed the paired associate technique for measuring learning ability, and have primarily investigated socioeconomic-status (SES) differences between intelligent test behavior and learning ability test performance. In each of the following studies, the high SES category consisted of white elementary school children while the low SES contained only black elementary children. The method used to rate the subjects was the Index of Status Characteristics, consisting of such data as the ranking of: occupation, source of income, house type, and dwelling area. It is important to note that the following studies did not present the paired associate task in accordance with a specific learning criterion (described previously by Reese and Lipsitt). Rather, these studies utilized the "study-test method," in which the subjects were given one study trial followed immediately by one test trial, for a total of n study-test sessions. Learning proficiency was, therefore, equal to the number of accurately recalled paired items within the allotted study-test sessions.

Semler and Iscoe (1963) and Green (1969) investigated the performance differences of high and low SES children on a paired associate learning task (noun pairs) with the Weschler Intelligence Scale for Children and the Lorge-Thorndike IQ, respectively. Similarly, Jensen (1968) compared high and low SES children's performance differences on the Raven Progressive Matrices with the digit-span learning task. The results of these studies demonstrated that although substantial race differences were indicated by the children's performance on either of the intelligence tests or the Raven Progressive Matrices, no such differences were observed on either the paired associate or the sequential (digit-span) learning tasks. Additional support for these findings was provided by Rohwer, Ammon, Suzuki, and Levin (1971). This study investigated high and low SES children in kindergarten, first and third grades. The PPVT, Children's Progressive Matrices (CPM), and paired associate learning task (four paired associate lists comprised of 25 noun pairs) were administered to each sample population. The results again showed that the differences in performance between black and white samples were large on both the PPVT and the CPM, while the paired associate task showed small differences between these populations. Additionally, these differences decreased significantly at the older grade levels, suggesting a diminishing deficit.

Many of the above studies (i.e., Thompson and Witryol, 1942; English and Kidder, 1969; Blue, 1970; Walther, 1969; and Carver and Dubois, 1967) have indicated a positive relationship between learning ability test performance and intelligence test behavior. Although these results may be expected, particularly since both measures tap some phase of learning (current or previously learned behavior, respectively), studies such as Semler and Iscoe, 1963; Green, 1963; Jensen, 1968; and Rohwer et al., 1971, stressed the significance of measuring ongoing behavior change. The latter studies suggest that the most accurate measurement of the "ability to learn" might be least effected by cultural differences.

Learning Ability Test Performance, Academic Achievement, and Intelligent Test Behavior

Although relatively few studies have investigated the possible relationship between children's performance on tests of learning ability and academic achievement, much of the research dealing with these two variables usually includes a third dimension--intelligence test behavior.

Wallen (1959) taught 283 high school students to associate a letter of the alphabet to each of four tones (155-cps; 470-cps; 1000-cps; and 1970-cps). Intelligence and academic test batteries were also given to each student. Statistical analysis of their performances demonstrated that there was a significant correlation between the speed of learning on an auditory discrimination task, academic achievement, and intelligence test behavior. Further support for these findings was offered by Stakes (1961) who investigated the relationship of 12 learning tasks, an achievement test battery and intelligence test behavior. This study employed 240 Caucasian and Black school children within the southeast portion of the United States. The procedures were varied between tasks in order to maintain subject interest. The results demonstrated a significant relationship between children's performance on various learning tasks and their performance on intelligence and achievement tests. Meuller (1968) investigated the relationship of six psychometric instruments (i.e., Stanford-Binet Intelligence Scale, Illinois Test of Psycholinguistic Abilities, Peabody Picture Vocabulary Test, Pictorial Test of Intelligence, The Coloured Progressive Matrices, and the Primary Mental Abilities Test) with two achievement tests (i.e., Wide Range Achievement Test and the New York Achievement Test) and four tasks requiring the subject to learn within the testing session (i.e., coding subtest from the WISC, Modified Prognostic Reading Test, a paired associate task, and a simple alternation task). This study employed a sample of 101 educable retarded children ranging from six to ten years of age. A mean-standard score of the four learning tasks was calculated and compared with the mean of the achievement test scores and the scores derived from the published norms of the six psychometric instruments. The results demonstrated that the composite score of the four learning tasks predicted achievement test scores as well as any of the psychometric instruments. Meuller concluded:

> . . . the overall results of this investigation suggest that tasks such as these, requiring the subject to demonstrate his ability to learn within the test situation, show considerable promise as predictors of academic achievement [p. 144].

Bonfield (1969) conducted a study to investigate the hypothesis that overall achievement of young institutionalized EMR's, IQs ranging from 50-75, would be predicted with higher validity by a regression equation based on a combination of the auditory-vocal association and auditory-vocal automatic subtests from the ITPA, the picture vocabulary subtest from the Picture Test of Intelligence, and the number facility and perceptual speed subtests from the Primary Mental Abilities Test for all age levels than by the coding (WISC), a paired associate task, and a sequential learning task combined. Although the results supported this hypothesis, the former combination of subtests did not adequately predict student's performance on individual reading or arithmetic criterion. Green and Rohwer (1971) designed a study to investigate the relationship between learning tasks (paired associate-noun pairs, and sequential learning tasks--digit span) academic achievement (teacher assigned grades in reading and arithmetic and a total reading score on the Stanford Achievement Test) and intelligence test behavior (Lorge-Thorndike Intelligence Test and CPM) to differences in socioeconomic-status. This study employed 60 elementary school children and designated them as either low, lower-middle or middle SES, according to the Index of Status Characteristics. The results showed that: 1) the sequential learning task, CPM, and the Lorge-Thorndike produced similarly significant correlations to measures of achievement, and 2) each of these measures demonstrated large performance differences between the SES levels. The paired associate task, however, appeared to be tapping an ability relatively independent of the other measures since it produced non-significant correlations with the other measures and all SES groups performed equivalently on this task.

The previous studies are highly representative of research efforts in this area since: 1) the findings generally demonstrated a positive relationship between intelligence test behavior, learning ability measures, and academic performance; and 2) few studies have utilized learning ability or academic assessment devices which measure performance to a specific learning criterion. Rather, learning task scores were calculated according to the number of correct responses on n (pre-determined) number of study-test sessions and the academic scores were derived from standardized achievement tests (measuring only what has been learned, whether taught or not). The following two studies utilized a learning test which trained a task to a specific learning criterion (95-100% accuracy over two consecutive trials). The number of trials and errors to criterion were used as the indicators of learning proficiency. However, similar to previous research, both studies used standardized achievement tests as the measurement devices of academic performance.

Black Box Test of Learning Ability, Academic Achievement, and Intelligent Test Behavior

Two studies, Campanella (1972) and Meeks (1973) were designed to investigate the usefulness of the Black Box Test of Learning Ability as a culture-free device for assessing children's learning potential.

Campanella (1972) employed 36 American Indian and 36 Caucasian children from 10 years 6 months to 14 years

9 months of age (fifth through seventh grades). The subjects were assigned to one of three experimental groups, each consisting of 12 children from both cultures. All subjects received a paired associate and sequential learning task (incorporated within the Black Box Test), as well as intelligence (WISC and the Quick Test) and achievement (Wide Range Achievement Test) tests. Each experimental group was assigned a different reinforcement procedure as follows: G1, consisted of both food and verbal reinforcement; G2, verbal reinforcement; and G3, no food or verbal reinforcement. The results of this study demonstrated that intelligent test behavior, achievement scores, and "unreinforced" performance on the Black Box Test, were significantly higher for the Caucasian sample population. However, the results further demonstrated that the Indian children slightly surpassed the Caucasian sample when systematic verbal and food reinforcement (G1) was provided to both groups. A systematic cross-cultural replication of Campanella's study was performed by Meeks (1973). This research utilized the multi-racial population within Hawaii, consisting of six major cultural groups: Japanese, Caucasian, Hawaiian, Chinese, Filipino, and Samoan. One hundred and ninety-six subjects (ranging from fifth to

seventh grades) were selected from two religiously affiliated and two public elementary schools. The Black Box Test of Learning Ability (paired associate and sequential learning tasks), the Quick Test of Intelligence, and the Wide Range Achievement Test were administered to all subjects. The author designated the Chinese, Japanese, and Caucasian groups as "advantaged" cultures, in contrast to the Hawaiian, Filipino, and Samoan cultures, which were considered "disadvantaged." The major findings of this study showed that while the WRAT and QT tests clearly separated the six races into two groups that coincided with the "advantaged" and "disadvantaged" categories hypothesized, performance on the Black Box Test remained stable over all six racial groups.

The studies of Campanella and Meeks are of major importance to the literature of learning ability since they: 1) demonstrate that learning a novel task to a specific learning criterion (via paired associate and sequential learning tasks) produces results which equates performance across several cultures, and 2) suggest the use of a specifically available device (Black Box Test of Learning Ability) for the assessment of learning potential.

Learning Ability Test Performance and Performance Tasks

An indepth study to investigate the relationship between children's performance on learning and performance tasks was designed by Friedrichs, Hertz, Moynahan, Simpson, Arnold, Christy, Cooper, and Stevenson (1971). Twenty-six girls and 24 boys ranging in age from 3.6 to 5.8 years were chosen from white middle and upper-middle-class families to participate in this study. The learning tasks included paired associate, serial, oddity, and incidental learning as well as category sorting, concept formation, and problem solving (provided on two levels of difficulty). The score for each learning task equalled the number of correct responses emitted within a preassigned number of test trials. The performance tasks consisted of: social imitation, impulsivity, reactivity, variability, following instructions, attention, persistence and level of aspiration. Specific experimental procedures were designed for each performance task and individual scores were obtained from each. The results demonstrated that although there was a high degree of differentiation within both the learning and performance tasks, some relevant intercorrelations between tasks did exist. Specifically, the attention task was significantly related to the learning scores on: oddity learning, observational learning, problem solving, and category sorting, while the following instruction performance task was significantly related to paired associate learning, serial memory, observational learning, incidental learning, and problem solving. The author concluded:

> . . . the ability to carry out instructions and pay attention appear to be important attributes for effective learning at the preschool level [p. 170].

Stevenson, Williams, and Coleman (1971) employed 50 four and five year-old "disadvantaged" children attending day care centers, in a study designed as a basic replication of Friedrichs _et al_. (1971). The results paralleled those of Friedrichs _et al_., even though diverse sample populations were emphasized.

Summary of the Reviewed Literature

The present review of the literature demonstrates that previous research has investigated the relationship between children's performance on learning tasks with performance in other areas such as intelligence test behavior, academic achievement, and various performance tasks. Only a few of these studies investigating the relationship between learning and academic achievement (Campanella, 1972; and Meeks, 1973) have employed learning tasks (Black Box Test of Learning Ability) which specify completion of the task contingent upon an accuracy criterion rather than a pre-established number of trials. Although many studies demonstrated that children's performance on learning tasks was typically correlated with academic achievement scores (measured by such techniques as: the Wide Range Achievement Test, New York Achievement Test, teacher ratings or grades) no studies have researched the relationship between learning tasks and specifically trained academic skills. It is noted that those studies which reviewed children's performance on learning tasks and its relation to school performance, typically disregarded the possible value of the learning curve. It was, therefore, the intent of this study to expand the research on the Black Box Test of Learning Ability by: 1) investigating the relationship between this device and children's performance on each of two academic skill areas (multiplication and writing), and 2) providing learning curve comparisons for each of the subjects across all tasks (via a single subject design and the use of a similar data collection system across all tasks).

METHOD

This study was conducted within the Pocatello School District (Number 25) during the second semester of the 1973-74 school year. Permission for this study was granted by those administrative personnel responsible for Tendoy Elementary School.

Selection of the Subjects

Tendoy Elementary School was randomly chosen from a selected population of five elementary school services by the Laboratory for the Diagnosis and Treatment of Learning Disabilities, an agency within the Pocatello School District, for participation in this study. Previous selection of the school tasks, according to various prerequisite criteria, limited the population to second grade students. Although two second grade classrooms were functional within Tendoy Elementary, the principal identified a specific preference for the selected classroom. All 24 students within this classroom were administered the multiplication and writing pre-tests. The most commonly errored letters were then identified (eight letters). Those subjects demonstrating 0% accuracy on both the writing (all eight letters incorrectly reproduced) and multiplication facts (errors on all twelve facts) were then selected as participants for this study. Twelve subjects qualified for this study, five boys and seven girls, and ranged in age from seven years three months to eight years two months.

Facility

The administration of this study was conducted in a two-room trailer, positioned approximately fifty yards from the building facility of Tendoy Elementary. The room designated for use was 12' by 20' and was divided into two sections by a paneling insert. Chairs and tables were provided in each section and the experimenter was free to use either space. Quiet and comfort were enhanced by carpeting throughout both sections. All portions of this study were conducted within this facility.

Selection of the School Tasks

The intent of this study determined that two school tasks be identified which would provide information concerning the Black Box Test as an indicator of individual learning ability, particularly as it relates to the educational setting. The selection of each task was contingent upon the following prerequisite criteria:

1. The selected subjects had to demonstrate naivete (0% accuracy on both pre-tests) for the chosen tasks. If the learning curves and scores were to be comparable to the novel tasks within the Black Box Test, the design had to also include comparable school tasks which were also novel to each of the subjects. This insured that all subjects would begin at the same level of competency. Because this requirement was difficult to fulfill at the higher grade levels, task selection within the second and third grades was emphasized.

2. Both tasks should have been of the same level of difficulty (i.e., traditionally taught to students of approximately the same age). Tasks taught at different grade levels would produce different learning curves among similarly aged students simply because one task would be more difficult than another.

3. The tasks should involve somewhat different skills, one of which must exemplify merely associative level learning. The other task should be representative of some other school-related skill (e.g., fine-motor coordination) as well. Associative-type learning is emphasized strongly at the lower grade levels.

4. The tasks should be representative in that they must involve frequently utilized skills within the basic curriculum of the selected grade level.

5. The tasks must be traditionally taught within the standard curriculum for the selected grade level. This prerequisite not only increases the relevance but also the novelty of this study.

It was, therefore, determined that: 1) learning the facts of a randomly selected multiplication table and 2) learning to accurately reproduce eight selected letters of the alphabet, cursive style, would satisfy each of the above task selection criteria.

In addition to the above, a number of prerequisite criteria were established to determine the specific method of presentation for these tasks:

1. The materials must be instructional in nature, thereby teaching the tasks to a specific accuracy criterion. Learning curves cannot be accurately provided by simply assessing whether or not the task was previously learned.

2. The school task presentations should follow the basic procedures of specific, commercially produced, methodologies and materials. This criterion is designed to limit experimenter bias by providing specific teaching

techniques within each task and to increase the relevancy of the results by utilizing methods which are presently available and feasible to teachers.

3. Both sets of materials should preferably be produced by the same author in order to provide an overall consistency across the selection of the teaching strategies.

4. Both methods must produce data which are readily translated into learning curves. Fulfillment of this requirement is necessary to establish legitimate comparisons between the individual learning styles of the Black Box Test and the school tasks. Learning curves from the Black Box Test are most meaningfully derived via the number of errors to a specific accuracy criterion.

5. Finally, the methods and materials must present objective formats for data collection. Continual monitoring of individual responses, rather than global achievement scores are necessary for the accurate production of learning curve data.

In order to meet each of the above prerequisites for method selection, it was determined that: 1) Hofmeister's Audio-Tutorial Math Program (Research Edition, 1972), and 2) Hofmeister's format for "Teaching Writing" (1973) would serve as the basic methodologies for teaching the multiplication and cursive writing tasks, respectively.

Description of the Black Box Test

The device utilized to provide an estimate (composite scores as well as learning curves) of children's "learning ability" was the Black Box Test of Learning Ability, which was compiled by Campanella (1972). The test consists of a paired associate task and a sequential learning task.

The test was housed in a black, hard-plastic enclosure, 6-3/4 inches wide and 11 inches long. The lid of the testing kit could be placed in a perpendicular position to provide a screen for various stimuli arrangements throughout the testing situation. The platform of the Black Box Test was structured as follows (proceeding from front to back): 1) Row A was located 1/2 inch below Row B, and contains five, 2-inch square slots (each separated by three-sided frames to prevent the stimuli from moving); and 2) a raised level containing, a) Row B, five slots aligned with the spaces on Row A in order to facilitate the matching tasks, and b) Row C, an extended flat-surface, three inches immediately behind Row B, allowing a space for stimulus display and block manipulation (Figure 1).

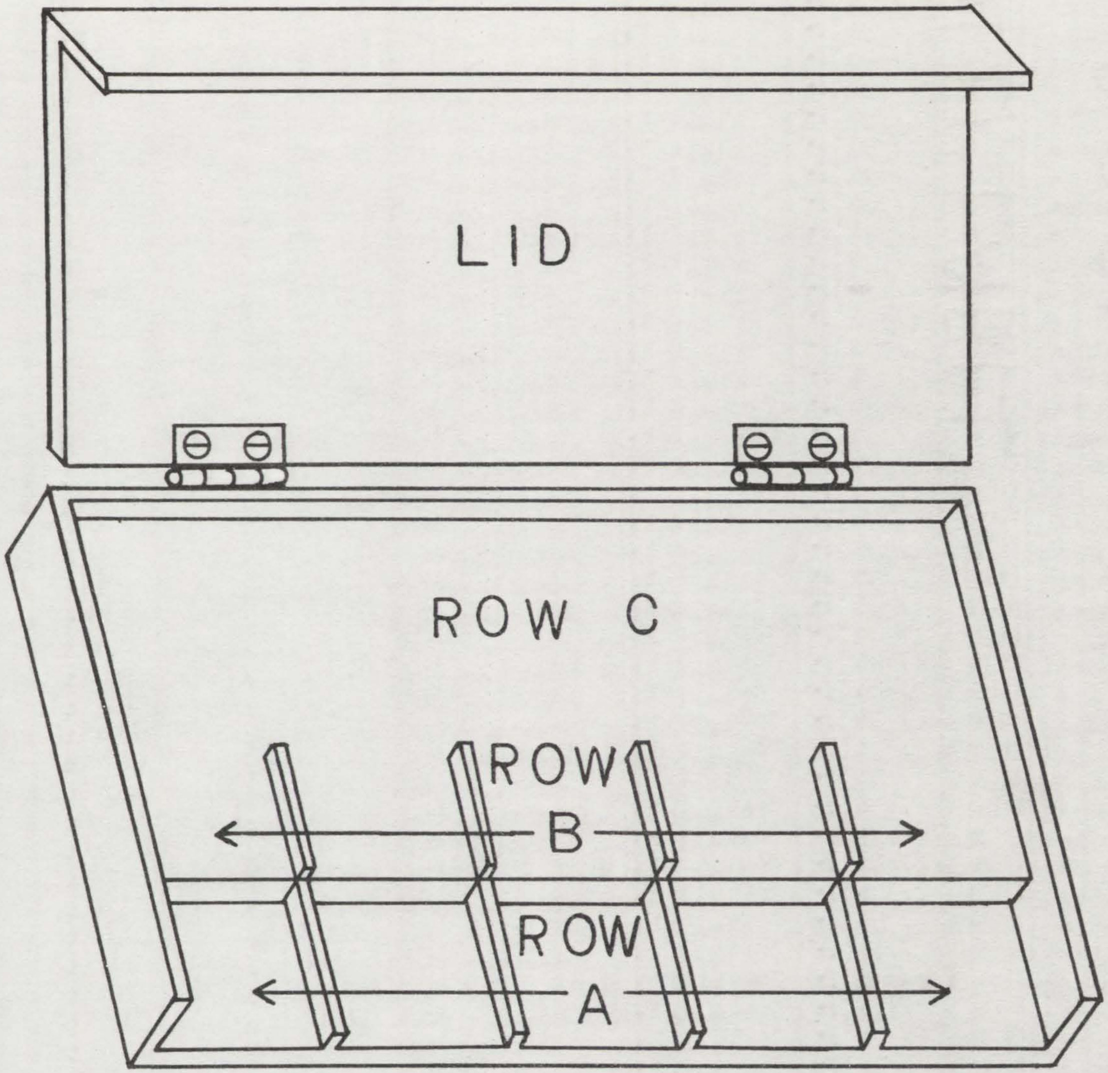


Fig. 1. Representation of the Black Box Test housing unit.

I. The Paired Associate Task

Administration. The stimuli used for the paired associate task were basic adaptations of the random shapes and dot patterns previously provided by Vanderplas and Garvin (1959) and Garner (1966) respectively. These stimuli and their specific pairings are random in nature and were devised to be equally unfamiliar to all subjects. The stimuli and the correct match between the "dot" and "pattern" blocks are represented in Figure 2. Each "dot" block on the top row corresponds to the "pattern" block immediately below in the following manner: A to 1; B to 2; C to 3, and D to 4. These pairings remained as "correct" matches throughout the entire paired associate task. Each of the blocks used in this learning task was made of hard-plastic and was 1-3/4 inches square by 1/4 inch thick.

This task required the children to learn four stimulus pairs (dot and pattern block associations) within the testing situation. There was no particular logic behind the various pairings and, therefore, each of the subjects had to develop his own method for learning. The specific instructions for the paired associate task are presented in Appendix A.

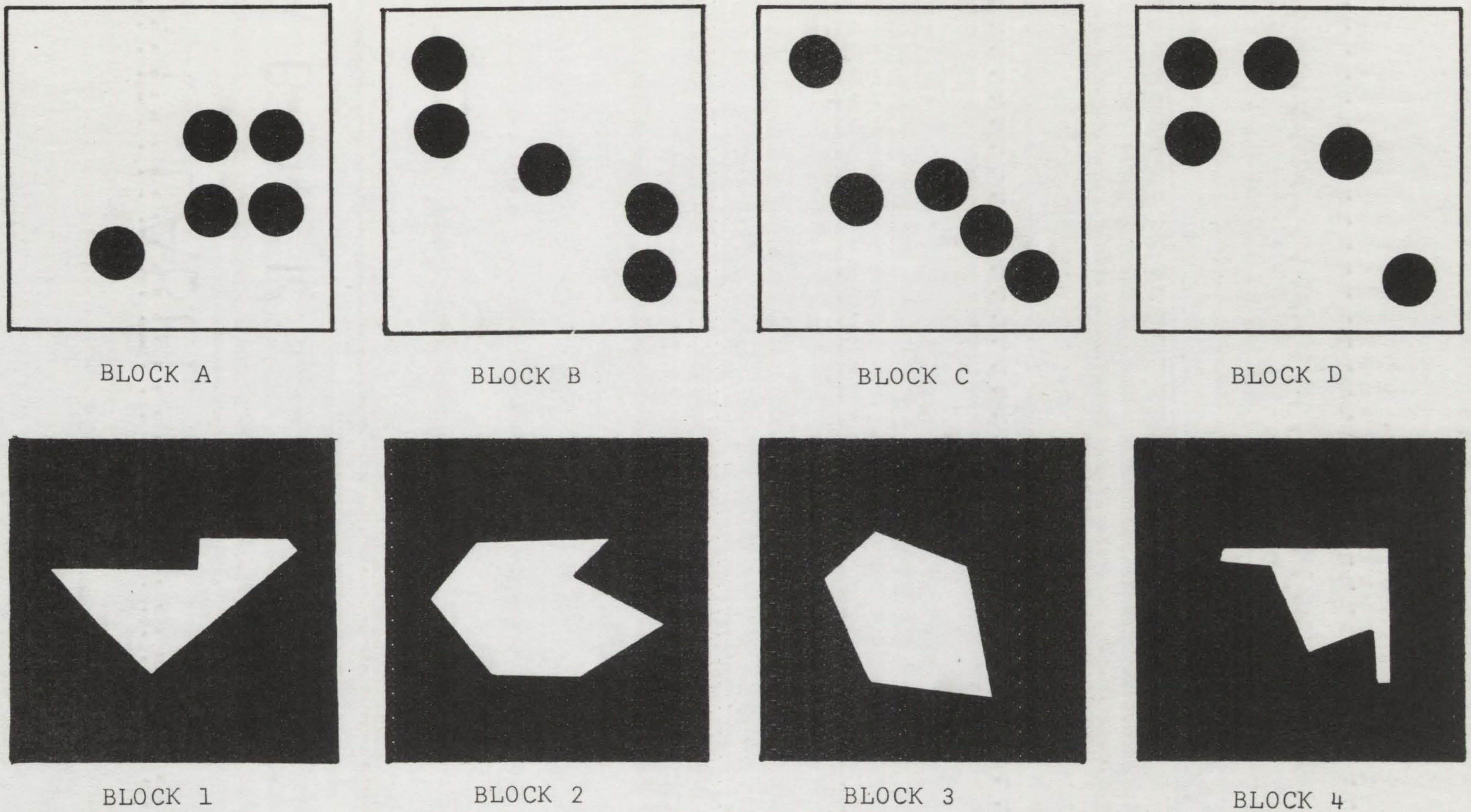


Fig. 2. Each of the plastic (1-3/4 inches square by 1/4 inch thick) stimuli used in the paired associate task, including the random dot patterns (Blocks A through D) and the random shape patterns (Blocks 1 through 4), are represented.

Initially, the "dot" blocks (labelled in Figure 2) were placed in order (A, B, C, and D) equidistantly across Row C on the Black Box platform (Figure 1). The "pattern" blocks were likewise placed in their respective slots (1, 2, 3, and 4 from left to right) in Row A corresponding to each of their correct matches in Row C. The subject was then shown how to correctly pair each of the stimuli by moving the blocks in Row C to the "match" in Row B. The subject was allowed to study the paired stimuli for four seconds, after which both rows of blocks were shuffled and relocated randomly in each of their respective rows. The pattern blocks were always located in Row A while the dotted blocks utilized Rows B and C. The examiner repeated the pairing procedure (moving the "dot" blocks from Row C to Row B), demonstrating that each trial would present both sets of blocks in different positions. The stimuli remained in their respective matched positions for a final study-interval of ten seconds. The lid was then raised to screen the tray and both sets of blocks were rearranged according to the randomized order presented at the top of column 1 on the paired associate score sheet (Appendix B).

A trial consisted of the correct pairing of each of the four stimulus pairs within each stimulus rearrangement. The subject was allowed as many pairings as necessary to complete each trial. Immediate feedback was provided by the examiner as to the accuracy of each response. In cases of incorrect responses, the dotted block was replaced in Row C and the subject was instructed to "Try another one," while correct responses remained in their positions in Row B. The stimuli were arranged on each trial by a fixed random order which repeated itself after 22 trials. After each trial, the lid was raised, and the examiner rearranged the blocks according to the assigned random order for that trial (paired associate score sheet, Appendix B). Completion of this task was contingent upon either: 1) meeting the accuracy criterion (two consecutive errorless trials) or, 2) reaching a ceiling level, defined as the completion of 44 trials. The ceiling level was exactly two complete cycles of the randomized stimulus sequence.

Scoring.[2] The composite paired associate score was derived for each subject by determining the total

---

[2]The number of errors per trial to criterion was

number of errors (each incorrect pairing) to criterion and subtracting it from 200, thus obtaining a positive score. The number of errors per trial were recorded at the bottom of the paired associate score sheet for each subject.

II. The Sequential Learning Task

Administration. Only Row B (Figure 1) of the black box platform was used for this task. Seven translucent, white blocks, without any markings, 1-3/4 inches square and 1/4 inch thick, were placed on the lid of the kit. The subjects were asked to pick one of the seven blocks while the examiner selected a second block (both were used for tapping various sequences). The remaining five blocks were then arranged across Row B of the platform.

The task consisted of five tapping patterns (designated on the left side of the sequential score sheet, Appendix B) across the five blocks arranged in Row B. Using a point of the block, the examiner demonstrated the first pattern with an even rhythm of ½ second per tap. The subjects then attempted to reproduce this same pattern immediately following each demonstration. Incorrect patterns were followed by a repeated demonstration of the same

---

the basis for the derived scores and learning curves for each of the subjects on all tasks.

pattern by the examiner. The specific instructions for this learning task are presented in Appendix A.

Each attempt to reproduce the tapping sequence was recorded as one trial. Criterion for completion of each pattern was two consecutive errorless responses, and the completion of all five sequences was required to complete the entire learning task. A ceiling level of 30 trials per sequence was specified as the maximum number of attempts allowed.

Scoring. The composite sequential learning task score was derived for each subject by determining the total number of errors (each incorrectly emitted pattern) to criterion across all five tapping sequences and subtracting this figure from 200, therefore, obtaining a positive score. The number of errors per trial were added across all five patterns and indicated at the bottom of the subject's sequential score sheet. For example, trial 1 was equal to the combined number of errors across the first trials of all five sequences. This was determined separately for each subject.

The combined Black Box Test composite score was determined by adding the paired associate score and the sequential score of each subject and then dividing by two.

The individual learning curves were based upon the number of errors to criterion for the combined Black Box Test performance. Therefore, the number of errors per trial were combined by adding the subject's performance across the paired associate and sequential learning tasks, trial by trial. For example, the first trial of both the paired associate and sequential tasks were combined and plotted on the learning curve as Black Box Test, Trial 1.

Description of the Multiplication Task

Administration. The multiplication facts (6's) were selected randomly from a population of the basic multiplication combinations, with the exception of 0's, 1's, 2's, 5's, and 10's. The latter facts were viewed as tasks too simplistic for providing differential learning curves. The pre-test (Appendix C), including all facts within the six times table, was administered to each subject within the classroom to determine naivete (0% accuracy). The following instructions introduced the math pre-test:

> Print your full name on the top of this paper. It has some multiplication problems I would like all of you to try and answer. Your teacher has not taught these facts to you yet, but I would like to see how many you might know anyway. Do as many as you can. If you have any questions, raise your

> hand. You will have five minutes to work. All-right, begin now.

The selected subjects (those exhibiting the naivete criterion) were introduced to the multiplication task with the following instructions:

> Select a chair and listen carefully. In front of each of you is a pair of headphones which will guide you through the following tasks. First, I will give you a study guide, like this (holding an example). You will study the items, saying each item aloud with the tape recorder. As soon as the tape is finished, please give the study guides back to me. Do not mark on these sheets. I will then give you a test sheet and the same items, without the answers, will be presented by the tape recorder. Complete each item, but please do not go faster than the tape recorder and do not say the answers aloud. After the test, I will correct each of your papers. While waiting for your turn, listen to the music through the headphones. We will repeat this until you get them all correct two times, one right after the other. Allright, we will now begin.

Each student selected one of three chairs (three students participated in each math session) in front of a 4' diameter, circular table. Vision from the other subjects' papers was blocked by a portable three-way divider (each leaf was 1' high by 2' long) made of particle board, and placed on top of the table.

The tape recordings were provided by a cassette tape recorder. Four tapes (a study guide and three test tapes coinciding with test sheet forms A, B, and C) were prerecorded to provide a consistent administration of this task. The study guide tape presented each problem and answer at five second intervals. The instructions on this tape were as follows: "Listen carefully. Look at your study sheet. I would like you to read each problem aloud with me as I say the problem on the tape recorder. Ready? Here we go." The tapes corresponding to the test sheets presented only the problems and allowed a five second interval for written responses. The instructions on each of the test tapes were as follows: "Have your test sheet and your pencil ready. If you have to change an answer, just cross it out and write it again. Do not erase. OK. Here we go."

The study guide (printed sheets) consisted of all twelve facts for the six times table and presented them in an increasing numerical order. Both questions and answers were provided and underscoring emphasized each of the answers. An example of the study guide is presented in Appendix C.

Three different test sheets (A, B, and C) were developed which examine the basic facts in three separately determined randomized orders. The testing sequence (form presentation orders) were administered according to the following randomization: C - A - B - A - B - C - B - A - C.

Each test sheet presented the question and provided a space for the subjects to complete the answer. Examples of each test sheet form are presented in Appendix C.

A study guide period always immediately preceded every testing procedure. Each cycle of the study guide-test sheet administration was equivalent to one trial. Each student received three trials per day. Task completion was contingent upon two consecutive trials at 100% accuracy. A ceiling level was established at 60 trials.

Scoring. The number of errors per trial to criterion was recorded for each subject. A composite math score was derived by subtracting the total number of errors to criterion for each subject from 400, thus obtaining a positive score.

Description of the Writing Task

Administration. The Zaner-Bloser cursive style for writing was used as the basis for establishing a criteria guide for this task. The writing task consisted of eight letters (d, q, f, y, b, i, n, and g), and the "Hofmeister Method for Teaching Writing" was followed as the guide for determining the teaching format. All children within the second grade classroom were administered the writing pre-test (Appendix D), consisting of the 26 letters of the alphabet in the low-case form. Those letters most frequently reproduced incorrectly throughout the pre-test were identified as "instructional letters" for this study.

The pre-test was administered with the following instructions: (The writing sheets were passed out to each student face down.)

> Write your name on the back of the paper. Now turn the paper over. In this sheet, the alphabet is written in cursive style. Your teacher has not taught this style of writing to you yet, but I would like to see how many each of you can do anyway. I would like you to write each letter exactly as I have written it in the lines immediately below. You have ten minutes to work on this. If you have any questions, raise your hand. Now begin.

All criteria necessary for specifically identifying the accuracy of each letter were pre-established before this pre-test and are presented in Appendix D. The criteria regarding time, direction, and continuous motion were not factors in correcting these pre-test responses.

The subjects selected as participants in this study were introduced to the writing task with the following instructions:

> Select a pencil. We are going to learn to write a couple of letters. First I will make the letter, then you will trace it, just the way I made it. You are then to make a letter just like it on the line immediately below. This will be different

> than printing since you are not to lift your pencil from the page while making the letter. You will not be allowed to erase your mistakes. Instead, we will try the whole letter again. I will show you your mistakes after each letter that you write, and will put a star on all correct letters. We will repeat these steps until you get two letters, one right after the other, completely correct. Then we will try a new letter. Remember, try and make your letter just like mine. Allright, ready to begin?

"White line manuscript paper" (8-½ by 11 inches), typically supplied in the second grade, was divided into thirds and used as the writing surface for each of the subject's responses.

In summary, the task involved the following sequence: 1) examiner produced the letter, 2) the subject traced the letter, 3) the subject produced the letter immediately below the model letter, and 4) the examiner corrected the letter. This sequence was repeated until the subject reached criterion for each letter (two consecutive errorless reproductions) or until the ceiling level was reached (60 trials per letter).

Scoring. Each incorrect trial was recorded as one error. The number of errors to criterion per letter was recorded for each subject. Learning curve data were provided by adding the number of errors across all eight letters, trial by trial, to criterion for each subject.

Therefore, the errors on the first trials of all letters were combined and plotted on the learning curve as Trial 1. Composite writing scores were calculated by subtracting the total number of errors for all eight letters per subject from 500 (to obtain a positive score).

A chart summarizing the previous discussion concerning each of the task's relation to: trial definitions, task criteria, composite score formulas, methods of collecting data for the learning curves, and ceiling levels are presented in Appendix E.

Teacher Rating

The teacher was asked to rank each of the participants within this study according to their overall academic achievement, in order to investigate the relationship between the Black Box Test, school tasks, and the teacher's evaluation for each of these students. She was instructed accordingly:

> Please rank each student within your classroom according to your impression of their overall academic performance. Try not to include your evaluation of their respective ability to learn or suspected intelligence in this ranking. Rather, simply rate them according to their averaged output, considering all school tasks such as reading, spelling, math, printing, etc. Rank the highest performer with a score of 1.

General Procedure

After each of the pre-tests were administered, 12 subjects were identified and divided into two groups of six subjects each. All subjects were then individually administered both learning tasks within the Black Box Test of Learning Ability. The administration of this device averaged 45 minutes and was always completed within one sitting. Subsequently, Group I (subjects labelled S1 - S6) received the writing followed by the math program, while Group II (subjects labelled S7 - S12) received the math program followed by the writing program. This procedure was designed to control for any differences due to sequence effects across programs. The subjects were brought individually to the trailer for letter training sessions. Groups of three, however, were employed in each math session and through the use of headphones, participated in the task simultaneously. Since only a limited amount of time per day was available (1-½ hours), only a few subjects could be trained daily. In order to limit the extraneous variable of forgetting, each subject, initiating a task, participated daily until reaching criterion. Subsequent subjects (within each group) were introduced as vacancies occurred. The daily training sessions were divided between writing and math tasks. Therefore, members from both Groups I and II participated each day. All subjects completed criteria on all tasks within a five and one-half month period.

RESULTS

Comparisons of the subjects' performance across all three tasks (Black Box Test, math, and writing) were investigated via two basic forms of analyses. Firstly, the total composite scores were derived for each subject across all tasks to: 1) provide the basis for a statistical analysis (Rank Difference Correlation), comparing the relationship between tasks, and 2) provide Z and T score comparisons for each of the subjects combined Black Box Test performance and their respective scores on the math and writing tasks. Secondly, the learning curves for each subject (trial by trial) across each task was provided and comparisons were made with the average learning curves calculated from the performances of all 12 participants. Learning curve analysis contained intra-subject comparisons of: 1) individual performance (progression of errors across tasks), and 2) length of the curves (number of trials to criterion across tasks).

Composite Score Analysis

The number of errors to criteria was the basis for calculating the composite scores for each of the tasks.

Each subject's error score for each task was converted into a positive score (composite score). The conversion formulas for each task are outlined in Appendix E. Each subject was ranked according to his composite score (rank score of 1 was assigned to the subject with the highest composite score). The teacher rating for each subject's overall academic performance could not be calculated into composite scores but was available as ranking data. Rank difference correlations[3] were calculated between: 1) each subtest within the Black Box Test, 2) the combined Black Box Test score, 3) math, 4) writing, and 5) teacher rating. Data providing the basis for these comparisons are presented in Appendix F. The rho results for each of these comparisons are presented in Table 1. Although a minimum alpha level of .05 was required for significance, higher level correlations are additionally designated.

The results (Table 1) indicate that although both subtests of the Black Box Test (the paired associate and sequential learning tasks) correlated significantly at the same level (.05) with math and writing, the highest rho

---

[3]The rationale for using the Rank Difference correlation is provided by Borg and Gall (1971). The number of subjects was less than 30 and ceiling limits for each task prevented collecting truly continuous data for all subjects.

Table 1

RHO Correlations Across Tasks and Teacher Rating

	Paired Associate	Sequential Task	Combined Black Box Test	Math	Writing	Teacher Rating
Paired Associate	1.00	.461	.940***	.706*	.699*	.392
Sequential		1.00	.659*	.589*	.596*	.320
Combined Black Box Test			1.00	.733**	.841***	.441
Math				1.00	.573	.633*
Writing					1.00	.615*
Teacher Rating						1.00

*Significant at the .05 alpha level (requiring rho greater than .576).
**Significant at the .01 alpha level (requiring rho greater than .708).
***Significant at the .001 alpha level (requiring rho greater than .823).

calculated values are represented consistently by the paired associate task. Furthermore, the combined performance on the Black Box Test (average of the composite scores of both subtests) exhibited correlations at higher levels of significance with both the math and writing tasks (.01 and .001 levels respectively) than either the paired associate or sequential learning tasks. It is also noted that only the school tasks were significantly correlated with teacher rating (.05 alpha level). Finally, the rho calculation between the math and writing task indicates that these tasks were not significantly correlated.

The composite scores for the combined Black Box Test (BBT) performance,[4] math, and writing tasks for each subject were translated into standard scores (Z scores) and are presented in Appendix F. Z score comparisons between the BBT and the math task and between the BBT and the writing task are plotted for all subjects in Figures 3 and 4 respectively. The Z score distributions have a mean of 50 and a standard deviation of 10. Larger standard scores indicate fewer errors and a higher composite score.

---

[4]All Z and T score calculations as well as the individual learning curves, representing Black Box Test performance, were based upon the combined performances of each of its subtests (paired associate and sequential learning tasks).

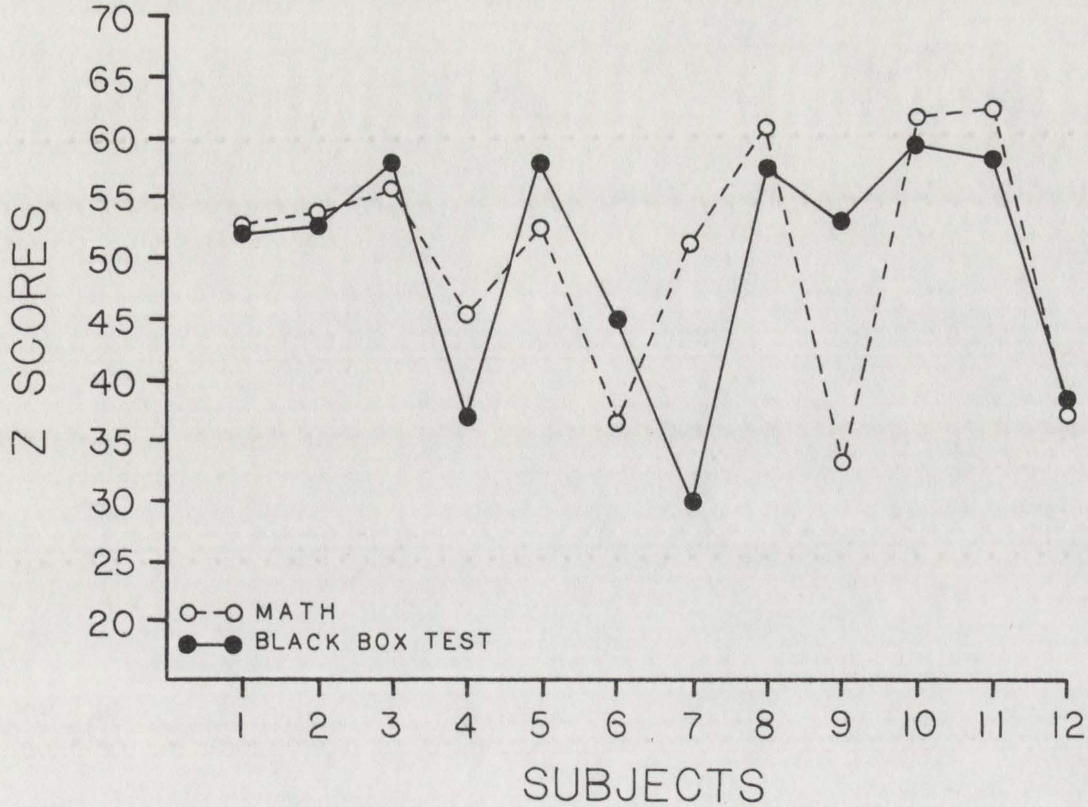


Fig. 3. Z score comparisons between Black Box Test (Combined) and Math for all twelve subjects.

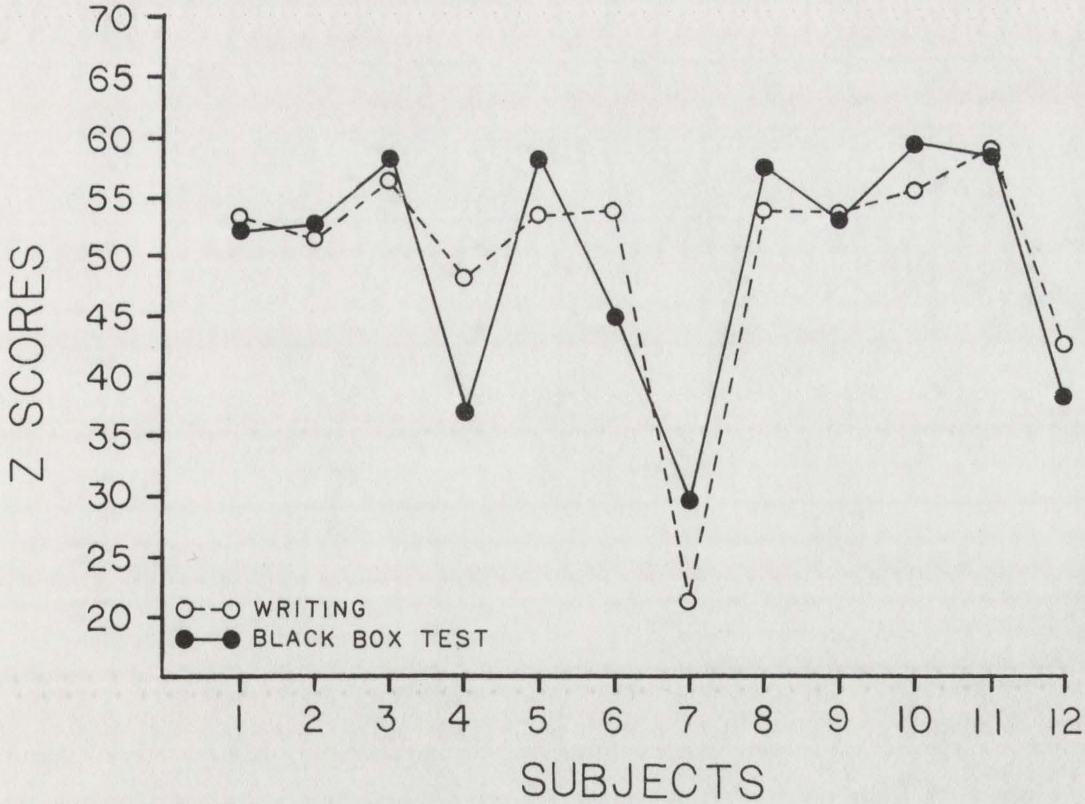


Fig. 4. Z score comparisons between the Black Box Test (Combined) and Writing for all twelve subjects.

Figure 3 indicates that, although a highly positive rho correlation (.733) exists between the BBT and the math task across all subjects, two of the subjects (Subjects 7 and 9) exhibited a difference larger than 1 standard deviation between these tasks. The stronger correlation (rho = .841) between the BBT and the writing task is clearly evident throughout Figure 4, particularly since only one subject (Subject 4) exhibited a writing score slightly more than 1 standard deviation from his respective Black Box Test performance.

Further investigation of Figures 3 and 4 suggests different population distributions across the various tasks. Specifically, the Black Box Test distribution indicates that nine of the 12 subjects (Subjects 1, 2, 3, 5, 6, 8, 9, 10, and 11) ranged within one standard deviation above or below the mean, while two subjects (Subjects 4 and 12) were two standard deviations and one subject (Subject 7) was three standard deviations below the mean. The math task distribution (Figure 3) placed six subjects within 1 standard deviation of the mean, three subjects (Subjects 8, 10, and 11) two standard deviations above the mean and three subjects (Subjects 4, 7, and 9) two standard deviations below the mean. Finally, all subjects on the writing task (Figure 4) performed within 1 standard deviation of the mean, with one exception (Subject 7), who was located three standard deviations below the mean. The math task, therefore, represents the most normal distribution, while the writing task grouped most subjects near the mean, discriminating one subject radically below the mean. The BBT distribution was negatively skewed, since none of the subjects was located beyond 1 standard deviation above the mean.

The composite scores for the combined Black Box Test, math and writing tasks were also converted into normalized standard scores (T scores) with a mean of 50 and a standard deviation of 10. The numerical data are presented in Appendix F and each subject's T scores for each task are compared in Figures 5 and 6.

The results in Figure 5 (T score comparisons between the BBT and the math task) were essentially equal to the subjects' Z score comparisons in Figure 3. Figure 6 (T score comparisons between the BBT and the writing task), however, indicates that Subject 4 obtained very similar T score comparisons, while Subject 6 produced T scores separated by more than 1 standard deviation. These results were directly reversed in Figure 4. Although there was a

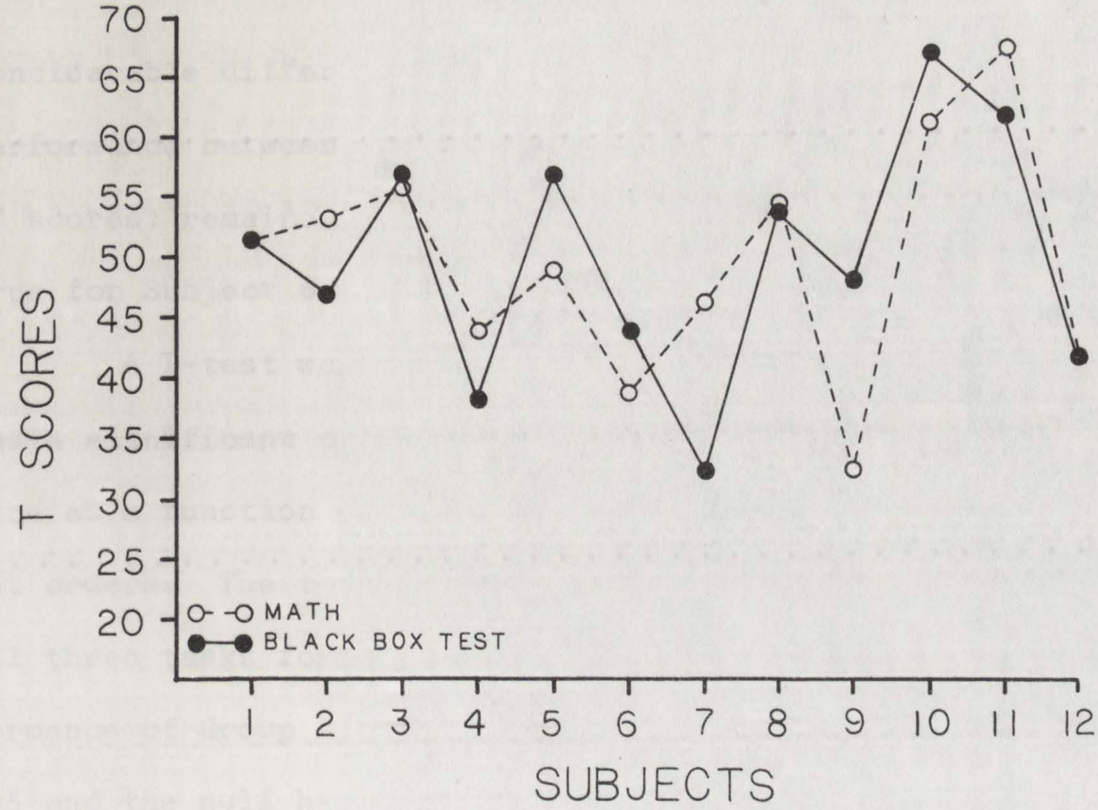


Fig. 5. T score comparisons between the Black Box Test (Combined) and Math, for all subjects.

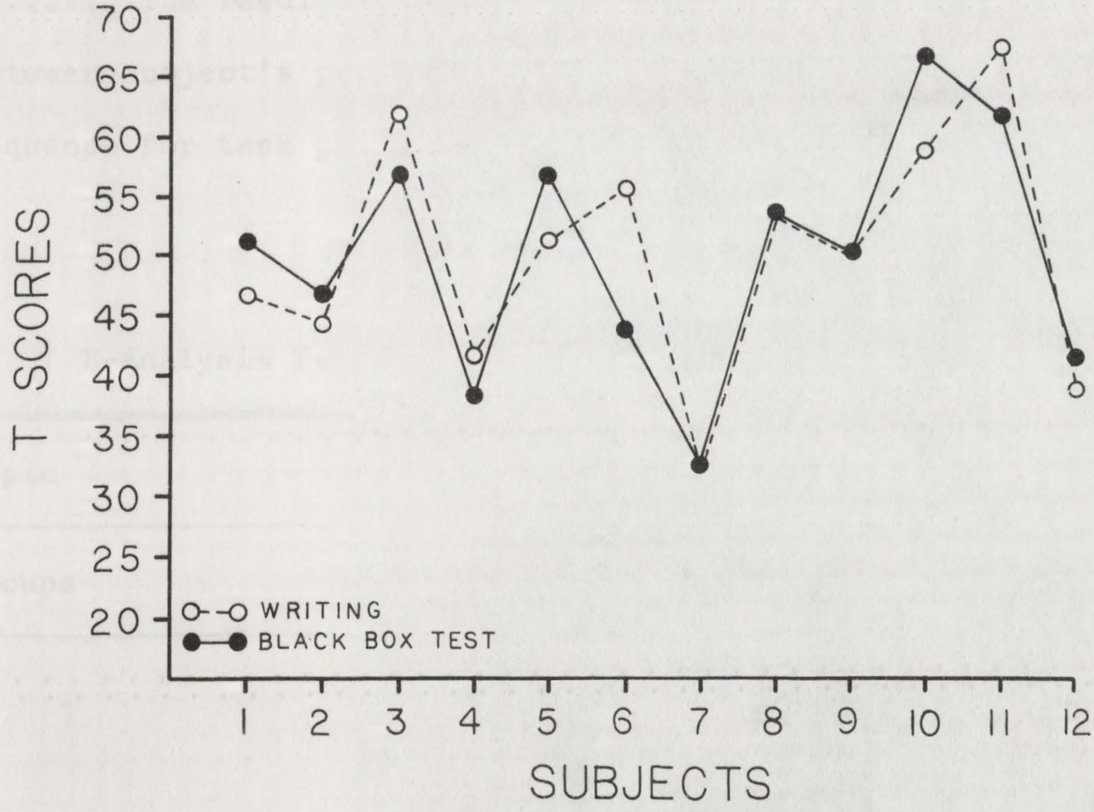


Fig. 6. T score comparisons between the Black Box Test (Combined) and Writing, for all subjects.

considerable difference exhibited by Subject 4's Z score performance between these tasks, the percentile ranking (T scores) remained relatively similar. The opposite was true for Subject 6.

A T-test was performed to determine whether there was a significant difference between the subjects' performance as a function of receiving the school tasks in different orders. The total number of errors to criteria across all three tasks for Group I was contrasted with the performance of Group II. The alpha level was established at .05 and the null hypothesis was reject4d if T -2.228 or +2.228. The results (Table 2) suggest that the differences between subject's performance was not due to a different sequence for task presentation.

Table 2

T-Analysis Testing the Difference Between Groups

Topic	df	T	P
Groups	10	-.0005	N.S.

Learning Curve Analysis

Analysis of each subject's individual learning curve investigated: 1) the performance levels (i.e., intrasubject comparisons of the progression of errors across all three tasks to the mean learning performance), and 2) the length of the curve (intrasubject comparisons of the number of trials to criterion across all three tasks to the mean learning curve length). The data for each of the subject's learning curves are presented in Appendix G. Appendix H (Subjects 1 through 12) presents the individual learning curves for all subjects and tasks. In order to lend additional information for the following learning curve comparisons, the proportion of trials within each standard deviation was determined, and the percentage distribution is presented in Appendix G. The relationship of the Black Box Test to each of the school tasks was indicated as similar only when the following criteria were met simultaneously: 1) the majority of trials for the compared tasks occurred on the same side of the mean learning curves, and 2) the largest percentage of trials (above or below the mean) between tasks was less than a 25% difference. Additional comparisons for subjects meeting these criteria were based upon individual learning curve similarities.

I. Performance Analysis

BBT and math task learning curve comparisons. The results indicated that seven of the 12 subjects (Subjects 2, 3, 4, 8, 10, 11, and 12) produced BBT learning curves similar to their respective math task performances. Specifically, for Subjects 2, 3, 8, 10, and 11, the majority of trials was below the mean (less errors per trial than the average), while Subjects 4 and 12 produced most responses (more errors per trial) above the mean.

For each of these seven subjects (with the exception of Subject 2), less than 10% difference was exhibited between the percentage distribution (above or below the mean) between tasks. Subject 2, however, exhibited 59% of the BBT trials, as compared with 81% of the math trials, below the mean learning curve (percentage distribution separation of 22%).

Further investigation revealed that only three of the seven subjects (Subjects 3, 10, and 12) exhibited highly similar learning curves between tasks (comparable beyond the mean analysis). Both learning curves of Subject 3 present a few initial trials above the mean, stabilizing most of the remaining trials within 1 standard deviation below the mean, and completing the task two standard deviations below the mean. Subject 10 produced comparably short learning curves consisting of trials closely aligned with the line indicating 1 standard deviation below the mean. Subject 12 illustrates, throughout both tasks, a comparable set of trials alternating beyond and within the region located 1 standard deviation above the mean. Although the other subjects (Subjects 2, 4, 8, and 11) demonstrated similar learning curves between tasks by comparing the percentage distribution across the means, analysis beyond this level was unreliable.

One subject (Subject 6) demonstrated relatively weak similarity between these two learning curves since 54% of the Black Box Test trials, compared with 90% of the math trials, ranged above the mean learning curves (separation of 36%).

The remaining four subjects (Subjects 1, 5, 7, and 9) did not produce similar BBT and math task learning curves. The BBT performances for Subjects 1, 5, and 9 show the majority of trials below the mean, while their respective math trials fall mostly above the mean. The opposite is true for Subject 7.

BBT and writing task learning curve comparisons. Ten of the 12 subjects (Subjects 1, 3, 4, 5, 7, 8, 9, 10, 11, and 12) demonstrated similar BBT and writing task learning curves. Six of these subjects (Subjects 3, 4, 5, 7, 11, and 12) produced learning curves with less than a 10% difference between the percentage distribution of trials above or below the mean, while four subjects (Subjects 1, 8, 9, and 10) indicated less than a 25% difference. Seven of these 10 subjects (Subjects 1, 3, 5, 8, 9, 10, and 11) produced learning curves on both tasks generally located below the mean learning curves. Subjects 4, 7, and 12 illustrated learning curves with most trials ranging above the mean.

Further investigation of these subjects reveals that seven subjects (Subjects 3, 5, 7, 8, 9, 10, and 12) produced highly similar learning curves between tasks (comparable beyond the mean analysis). Subject 3 exhibited a few initial trials above the mean (22% and 31% of the BBT and writing task respectively), followed by a response pattern ranging within 1 standard deviation below the mean, and terminating two standard deviations below the mean. Subjects 8 and 10 demonstrated similar learning curves since nearly all trials throughout both tasks ranged within the region 1 standard deviation below the mean. Subject 5 began below the mean, followed by a few trials (17% and 23% of the BBT and writing task, respectively) above the mean, and then stabilized within the region 1 standard deviation below the mean. The similarity for Subject 7 is clearly evident since both curves exhibited the majority of trials at least two standard deviations above the mean. Subject 9 initially responded within 1 standard deviation above the mean (30% and 36% of the BBT and writing trials, respectively) and terminated the majority of the remaining trials (68% and 52% of the BBT and writing trials, respectively) within 1 standard deviation below the mean. Finally, Subject 12 exhibited a pattern of responses throughout both tasks which alternated between 1 and 2 standard deviations above the mean.

Learning curve comparisons beyond the mean analysis did not produce highly similar results for Subjects 1, 4, and 11.

The two remaining subjects (Subjects 2 and 6) did not produce similar BBT and writing task learning curves. For Subject 2, the majority of BBT trials occurred below the mean learning curve, while the writing task produced most of the trials (71%) above the mean. The opposite pattern was true for Subject 6.

II. Learning Curve Length Analysis

The length of the BBT learning curve is the second aspect for investigating the usefulness of this device as an indicator of learning styles. The mean number of trials across all subjects for each task is indicated by the vertical line labelled "average number of trials" and connects the abscissa of each task in Figures 7 through 18. The standard deviations were calculated for each task and the number of trials for each subject across all three tasks were converted to standard scores (Z scores with a mean of 50 and a standard deviation of 10). These data are presented in Appendix G.

The length of the Black Box Test learning curve was considered positively related to the length of the school task learning curves if: 1) the standard scores were on the same side of the mean (both either above or below a standard score of 50), and 2) the distance between the standard scores was less than 1 standard deviation. The results (Appendix G) indicate that the length of the BBT learning curve was positively related to the length of the curves for nine subjects on both the math (Subjects 1, 2, 3, 4, 5, 6, 10, 11, and 12) and writing tasks (Subjects 2, 3, 4, 5, 6, 8, 10, 11, and 12). Only

Subject 9 produced learning curve lengths on both school tasks directly contrasting with his respective BBT learning curve length. Subjects 2, 3, 4, 5, 10, 11, and 12, however, produced BBT learning curve lengths positively related to both school tasks curve lengths.

Additional support for the above findings was provided by the Rank Difference Correlation (Table 3) calculated between the BBT learning curve lengths and each of the school tasks. A .05 alpha level was established as the minimum significance level.

Table 3

Rank Difference Correlation of Learning Curve Lengths Between the BBT and Each School Task

	Math	Writing
BBT	.741**	.801**

**Indicates significance at .01 alpha level (requiring rho > .708).

The results indicate that the length of the Black Box Test learning curve is significantly related to both the math and writing task learning curve lengths.

The learning curve length comparisons further demonstrated that the one subject (Subject 12), who reached the ceiling level on the BBT (sequential task), similarly reached the ceiling level on the writing task and exhibited one of the longest math task learning curves (surpassed only by Subject 9, requiring two additional trials beyond Subject 12). Subject 7, who required more trials than any other subject before reaching criteria on the BBT, was also terminated on the writing task at the ceiling level. Longer learning curve lengths exhibited on the BBT usually indicated similar learning curve lengths on the school tasks.

GENERAL DISCUSSION

Discussion of the Composite Score Analysis

The results of this study, based upon the composite score analysis, supports hypothesis 1. Each of the three Black Box Test scores, derived from the paired associate, sequential, and the combined BBT performance, significantly correlated with second grade children's scores on each of the school tasks (Table 1). This finding was expected, since all learning tasks within the BBT as well as both school tasks required a common skill, associative learning (i.e., learning through memorization rather than problem solving or concept formation). Specifically, the paired associate task involved the basic association of four stimulus pairings while the sequential task required the subjects to memorize five tapping sequences. The combined BBT performance was simply the average of the previous scores. The math task, requiring the subjects to memorize 12 basic multiplication facts, was essentially a paired associate task. The rho correlation between the math and paired associate learning task was .706, while the rho correlation decreased to .589 when compared with the sequential learning task. Writing, although involving another skill area (i.e., fine motor coordination) required the subjects to associate a number of stimuli (various segments of each letter) in a sequential order.

However, further investigation of the rank difference correlates (Table 1) reveals that the combined BBT performance correlates higher with school task performance than either individual BBT task. This may be partly because the combined BBT score provides a larger sample of associative learning.

Hypothesis 2, predicting that the combined BBT score would correlate higher to math than to the writing task, was not supported by this study. The results indicated a rho correlation of .733 (significant at the .01 level) for math as compared to the significantly higher writing task rho correlation of .841 (.001 level of significance). Furthermore, the standard score comparisons between tasks for each subject (Figures 3 through 6) reveal two of the 12 subjects (Subjects 7 and 9) exhibited a separation larger than 1 standard deviation between the math and Black Box Test scores, while only one subject (Subject 4) exhibited a writing score slightly larger than

1 standard deviation from his respective Black Box Test performance.

This hypothesis was based upon the research and postulates proposed by Jensen (1969) and Robb (1972). The former suggests that the majority of learning takes place on two levels: 1) Level I or associative learning, which basically requires the subjects to memorize either stimulus pairs or a stimulus sequence; and 2) Level II or problem solving learning, in which the subject is required to utilize previously learned stimulus sets in various strategies and combinations to solve novel problems. Robb (1972) discusses another major type of learning, effective learning. This level, although associated with cognitive learning, specifically relates to the motor output of an individual, requiring the learner to speak, write, move or perform some physical act to communicate what he has learned. Aligned with Jensen's proposal that the sequential and paired associate tasks are examples of associative learning and, in view of the fact that the math task merely required the association of 12 stimulus pairs, the highly significant correlation found between these tasks was expected. However, the stronger correlation between the Black Box Test and writing task, involving not only associative learning (sequential association of the various segments of each letter), but effective learning as well, was not predicted.

These findings, therefore, cannot be adequately explained solely on the basis of learning levels. A series of studies conducted by Sheffield and associates (Maccoby and Sheffield, 1961; Margolius, Sheffield, and Maccoby, 1961; and Weiss, Sheffield and Maccoby, 1961), and subsequently summarized by Sheffield (1961), however, suggest that strong consideration should be placed upon the variable of practice as a function of individual learning performance. Their results conclude:

> . . . attempted practice after a brief segment of the total demonstration will result in more accurate overt performance of that segment than attempted practice only at the end of the entire demonstration. That is, if practice came only at the end, the learner would forget a large part of what was shown by the end of a complete demonstration and would be unable to perform when called upon to perform [p. 63].

Furthermore, Michael and Maccoby (1961) designed a study to investigate the effect of "feedback" on learning performance of an adult audience viewing a factual film. The results demonstrated that learning performance was largely a function of providing knowledge of the correct response after practice. These findings applied to both overt and covert practice conditions as well as to both the more "intelligent" and less "intelligent" participants.

Analysis of the instructional units for the math and writing tasks revealed that the math format provided less appropriate practice, since the subjects were required to emit the written responses to the multiplication facts only after all 12 combinations had been presented in the study session. Although each of the multiplication facts was repeated orally throughout each study session by all subjects, requiring written responses instead would have assured more attention and a closer approximation to the terminal objective. Secondly, the writing program provided more "immediate feedback" than the math task format. Each attempted letter was corrected by the examiner before the subject was allowed to retrace the master letter. The math program, however, provided feedback only after the subject had attempted all 12 combinations. It appears that the correlations were highest between the writing task and the combined BBT performance, since both tasks within the BBT similarly emphasized "attempted practice" for small segments of each task, as well as "immediate feedback" following each response. It is proposed, therefore, that the method of instruction (the amount of attempted practice and immediate feedback) was the major factor in predicting the learning performance of individuals. This study specifically demonstrates that, although the combined BBT performance is highly related to school tasks requiring rote memory skills, the degree of this relationship is strongly related to the method of instruction. Increasing the amount of immediate feedback and attempted practice within the instructional format of school tasks increases the relationship of that learning performance to the performance exhibited on the Black Box Test of Learning Ability.

The rank difference correlation calculated between the math and writing tasks (Table 1) produced a non-significant rho value of .573. This result suggests that the math and writing tasks represented different school-related skills.

Although the students' performance on each school task correlated significantly (Table 1) to their overall school achievement as viewed by the teacher (teacher rating), their respective BBT scores did not coincide with the teacher's impression of their scholastic performance. Apparently, the school tasks were a relevant portion of the subjects' immediate school environment, while the BBT was further removed and less related to overall academic performance (as rated by the teacher).

Discussion of the Learning Curve Analysis

The individual learning curves demonstrated a stronger relationship between the learning styles associated with the BBT and the writing task, than the math task. Ten of the 12 subjects (83%) produced similar learning curves via mean analysis (determining whether the majority of responses ranged above or below the mean across both tasks) when the learning test and the writing task were compared. However, only seven of these subjects (70%) produced highly similar learning curve comparisons (beyond the mean analysis) by exhibiting highly comparable topographies. The learning curve comparisons between the BBT and math task revealed that only seven of the 12 subjects (58%) produced similar learning curves (via mean analysis) while only three of those seven subjects (43%) demonstrated highly similar topographies.

The learning curve length of the BBT was positively related to the lengths of the curves for nine subjects on both the math and writing tasks. The rank difference correlation similarly indicated a significant rho value between the BBT learning curve length and both school tasks (each at the .01 level of significance).

These results demonstrated that the most comparable individual learning curves were produced by those tasks (BBT and writing task) involving the most similar methods of instruction. Although the composite score analysis revealed a significant relationship between the BBT and the math task, the individual learning curves (trial by trial comparisons) were not highly related. The length of the learning curves (basically representing a total score) was the most reliable measure of the learning curve analysis.

Suggestions for Future Research

An analysis of the present study indicates that, although the BBT composite scores correlated significantly with children's performance on various school tasks, the relative educational usefulness of the individual learning curve was limited. While the length of the BBT learning curve generally served as an indicator of the length of the learning curves for each school task, the performance (trial by trial) comparisons between tasks were less reliable (particularly between the BBT and the math task). Further investigation of these results revealed that the major factor relating the BBT to each of these school tasks was the method of instruction rather than the learning test's assessment of particular "types" of learning ability.

In view of these findings, it is concluded that further efforts towards the standardization of the Black Box Test of Learning Ability, in its present form, would require more time and effort than could be justified by its potential usefulness for the educational setting.

The author suggests that a more functional method of assessing learning ability and producing relevant learning curves would be to obtain behavioral samples from the required materials and tasks rather than utilizing correlated activities. For example, students within a third grade math program could be administered pre-tests to determine their respective naivete levels. They could then be taught a series of novel tasks to a learning criterion from the program itself. Learning curves, as well as composite scores, could readily be obtained from this material. The major educational programs within a district curriculum (at all grade levels) could be standardized according to this technique in order to allow for intersubject comparisons and to increase the relevance of the individual data. This approach would maintain the major benefits previously discussed within the design of the Black Box Test (i.e., teaching the task within the testing session, establishing a learning criterion, using somewhat novel items, and producing learning curves), while increasing its relevancy to the educational setting. This approach to learning ability assessment would increase the probability of making immediately useful educational decisions since the assessment device and the educational program materials would be the same.

The above discussion is merely an expression of the author's preference concerning future "learning ability" research. It is recognized that researchers may wish to place continued emphasis upon the development of the Black Box Test. The following suggestions, therefore, are submitted in identification of those areas requiring further clarification and development, if the Black Box Test is to become a meaningful assessment of learning ability.

Initially, research should investigate the relationship of the BBT to school tasks representing each of the skill areas necessary for adequate school performance. Associative and problem solving learning (theorized by Jensen) as well as other required skill areas such as simple and complex discrimination, affective and effective learning are examples of such tasks. Such a study would primarily determine whether the BBT can be useful as an indicator of children's performance on school tasks which are not novel, but rather require the use of previously learned skills to demonstrate competency on the selected task (e.g., problem solving).

Secondly, further research concerning the method of instruction should be investigated to determine its role in effecting significant correlations between the BBT and school tasks. Apparently, increasing the similarity between such variables as the frequency of attempted practice and immediate feedback, increases the relationship between tasks, regardless of the type of learning involved. It would be interesting to determine, for example, whether the BBT performance (administered individually) correlates with students' learning novel tasks via group instruction.

The results of this study suggest that the relationship of the BBT to children's school performance may be limited merely to associative level tasks and, furthermore, may be useful only as an indicator of school tasks with a similar method of instruction. Such highly specific relationships would certainly limit the educational usefulness of the BBT. However, if the above suggested studies provided evidence indicating that the relationship between school tasks and this learning test was not limited to these specific variables, subsequent research efforts should provide:

1. a standard scoring procedure for each BBT learning task, as well as the calculation procedure of the combined performance score;

2. a standard procedure for producing learning curves (representing both the individual learning performance as well as the normative learning curves);

3. information concerning the variables within the test itself (e.g., the number of stimulus pairs, the length of the sequence patterns, complexity of the tasks, level of difficulty, etc.). The usefulness of this device for indicating learning styles of different age groups may necessitate a wider display of items and tasks; and

4. a standardization of the various BBT tasks across a multicultural population, with a minimum age range of six to twelve years. This is necessary if individual BBT scores are to be readily meaningful to educators utilizing this device.

SUMMARY

This study investigated the usefulness of the Black Box Test of Learning Ability as an indicator of children's performance (via individual learning curve and composite score analysis) across two selected school tasks. Twelve children, ranging in age from 7 years 3 months to 8 years 2 months, were selected from the second-grade classroom at Tendoy Elementary School (presently being serviced by the Learning Laboratory of Pocatello) to participate in this study. After demonstrating naivete on both a math and writing task, each subject was individually administered the Black Box Test of Learning Ability. The 12 subjects were divided into two groups of six subjects each. Each group received a different task presentation order, that is, Group I (Subjects 1 through 6) received the writing followed by the math task, while Group II (Subjects 7 through 12) received the math followed by the writing program. Composite scores were derived for all subjects and tasks (three BBT scores, math and the writing task). The results were discussed according to: 1) a statistical analysis of the composite scores, and 2) individual learning

curve comparisons, provided for all subjects across each task.

Although significant correlations existed between the three learning test composite scores and each school task, the combined BBT score produced a significantly higher correlate to math and writing than either the paired associate or sequential learning tasks alone. Stronger correlations were evident between the combined BBT composite score and writing than between the BBT and the math task. While the length of the BBT learning curve was indicative of the learning curve lengths for each school task, the performance (trial by trial) learning curve comparisons were less reliable (particularly between the BBT and the math task). The method of instruction, including the amount of attempted practice and immediate feedback, was suggested as the major factor correlating the Black Box Test to each school task.

In general, the composite score correlations were not demonstrably higher than various "ability" tests presently being utilized. Furthermore, the individual learning curve comparisons produced less reliable and oftentimes contradictory results and were, therefore, limited in their usefulness in making accurate educational decisions. In view of these results, it is concluded that further efforts towards the standardization of the Black Box of Learning Ability, in its present form, would require more time and effort than could be justified by its potential usefulness for the educational setting. Rather, it is suggested that future research concerning "learning ability" assessment emphasize techniques which would obtain a behavioral sample of the task itself, rather than utilize a correlated activity.

REFERENCES

Anastasi, Ann. Psychological testing. New York: Macmillan, 1961.

Becker, W. C., Englemann, S. & Thomas, D. Teaching: a course in applied psychology. Chicago, Palo Alto, Toronto, Sydnew: Science Research Associates, Inc., 1971.

Blue, M. C. Influence of mode of presentation, age and intelligence on paired-associate learning. American Journal of Mental Deficiency, 1970, 74(4), 527-532.

Bond, E. A. Tenth grade abilities and achievements. Teachers College Contributions to Education, 1940, No. 813.

Bonfield, J. R. Predictors of achievement for educable mentally retarded children. A Dissertation Abstracts International, 1969, 30(1-3), 1009A.

Borg, W. R. & Gall, M. D. Educational research, an introduction. (2nd ed.) New York: David McKay Company, Inc., 1971.

Campanella, S. Performance on learning tasks as a cross-cultural test of intelligent behavior. Unpublished doctoral dissertation, University of Montana, 1972.

Carver, R. P. & DeBois, H. The relationship between learning and intelligence. Journal of Educational Measurement, 1967, 4(3), 113-137.

Clelland, D. L. & Toussaint, I. The interrelationships of reading, listening, arithmetic computation and intelligence. Reading Teacher, 1962, 15, 228-231.

Drever, J. A. A dictionary of psychology. Baltimore: Penguin Books, 1965.

English, R. A. & Kidder, J. W. Note on the relationship among mental ability scores, teacher rankings and rate of acquisition for four year old kindergarten children. Psychological Reports, 1969, 24, 554.

Fahmy, M. Initial exploring of the shilluk intelligence. Cairo: Ear Misr Printing Co., 1954.

Friedricks, A. G., Hertz, T. W., Moynahan, E. D., Simpson, W. E., Arnold, M. R., Christy, M. D., Cooper, C. R. & Stevenson, H. W. Interrelations and learning and performance tasks at the preschool level. Developmental Psychology, 1971, 4(2), 164-172.

Garner, W. R. To perceive is to know. American Psychologist, 1966, 21, 11-19.

Gaw, F. A study of performance tests. British Journal of Psychology, 1925, 15, 374-392.

Goodenough, F. L. Mental testing: its history, principles, and applications. New York: Rinehart, 1949.

Green, R. B. SES differences on learning and ability tests in black children. Unpublished master's thesis, University of California, Berkeley, 1969.

Green, R. B. & Rohwer, W. D. SES differences on learning and ability tests in black children. American Educational Research Journal, 1971, 8, 601-609.

Havighurst, R. J. What are the cultural differences which may effect performance on intelligence tests? In K. Eells, A. Davis, R. J. Havighurst, V. E. Herrick, & R. Tayler. Intelligence and cultural differences. Chicago: University of Chicago Press, 1951.

Hofmeister, A. Audio-tutorial program, math. Instructional Technology Project, Exceptional Child Center, UMC-68, Utah State University, Logan, Research edition, 1972.

Hofmeister, A. M. Let's get it write. Teaching Exceptional Children, Fall, 1973, 30-33.

Hoveland, C. I. Human learning and retention. In S. S. Stevens (Ed.), Handbook of experimental psychology. New York: Wiley and Sons, 1951.

Jensen, A. R. IQ and scholastic achievement. Harvard Educational Review, 1969, 39(1), 111-123.

Jensen, A. R. Patterns of mental ability and socioeconomic status. Proceedings of the National Academy of Sciences, 1968, 60, 1330-1337.

Klineberg, O. Negro intelligence and selective migration. New York: Columbia University, 1935.

Lee, S. E. Negro intelligence and selective migration: a Philadelphia test of the Klineberg hypothesis. American Sociological Review, 1951, 16, 227-232.

Lehmann, I. J. Rural-urban differences in intelligence. Journal of Educational Research, 1959, 52, 62-68.

Maccoby, N. & Sheffield, F. D. Combining practice with demonstration in teaching complex sequences: summary and interpretation. In Student response on programmed instruction: symposium on programmed instruction. Chapter 5, National Academy of Sciences and National Research Council, Washington, D.C., 1961.

Margolius, G., Sheffield, F. D., & Maccoby, N. Repetitive versus consecutive demonstration and practice in the learning of a serial mechanical-assembly task. In Student response on programmed instruction: symposium on programmed instruction. Chapter 6, National Academy of Sciences and National Research Council, Washington, D.C., 1961.

McGeoch, J. A. The psychology of human learning. New York: Longmans, Green, and Co., 1942.

Meeks, B. A cross-cultural assessment of learning ability. Unpublished doctoral dissertation, Utah State University, 1973.

Michael, D. N. & Maccoby, N. Factors influencing the effects of student participation on verbal learning from films: motivating versus practice effects, "feedback," and overt versus covert responding. In Student response on programmed instruction: symposium on programmed instruction. Chapter 18, National Academy of Sciences and National Research Council, Washington, D.C., 1961.

Mitchell, M. B. The revised Stanford-Binet for university students. Journal of Educational Research, 1943, 36, 507-511.

Mueller, M. W. Validity of six tests of ability with educable mental retardates. Journal of School Psychology, 1968, 6(2), 136-145.

Mussen, H. P., Conger, J. J., & Kagan, J. Child development and personality. New York: Harper and Row, 1969.

Reese, H. W. & Lipsitt, L. P. Experimental child psychology. New York and London: Academic Press, 1970, 196-197.

Robb, M. D. The dynamics of motor-skill acquisition. Englewood Cliffs, N.J.: Prentice-Hall, 1972.

Rohwer, W. D. Learning, race and school success. Review of Educational Research, 1971, 41(3), 191-211.

Rohwer, W. D., Ammon, M., Suzuki, N., & Levin, J. R. Population differences and learning proficiency. Journal of Educational Psychology, 1971, 62(1), 1-14.

Rosandic, R. & Bukvic. A. Effects of the socio-cultural background upon success in solving tests of the Binet-Simon Scale. Psiholigija, 1970, 3(2), 217-221.

Semler, I. J. & Iscoe, I. Comparative and developmental study of the learning abilities of negro and white children under four conditions. Journal of Educational Psychology, 1963, 54, 38-54.

Sheffield, F. D. Theoretical considerations in the learning of complex sequential tasks from demonstrations and practice. In Student response on programmed instruction: symposium on programmed instruction. Chapter 2, National Academy of Sciences and National Research Council, Washington, D.C., 1961.

Sherman, M. & Key, C. B. The intelligence of isolated mountain children. Child Development, 1932, 3, 279-290.

Shrag, J. A. Idaho state: special education needs assessment study, November 1973.

Stake, R. E. Learning parameters, aptitudes and achievements. Psychometric Monographs, 1961, 6.

Stevenson, H. W., Williams, A. M., & Coleman, E. Interrelations among learning and performance tasks in disadvantaged children. Journal of Educational Psychology, 1971, 62(3), 179-184.

Thompson, G. G. & Witryol, L. The relationship between intelligence and motor learning ability. Journal of Psychology, 1946, 22, 237-246.

Vanderplas, J. M. & Garvin, E. A. The association value of random shapes. Journal of Experimental Psychology, 1959, 57, 147-154.

Wallen, N. E. A group test of auditory discrimination learning and some correlates with intellectual performance. The Journal of Psychology, 1959, 48, 263-270.

Walther, C. J. An investigation of paired associate learning within and across intelligence levels and sense modalities. A Dissertation Abstracts International, 1970, 30(11-12), 4842-4843.

Weiss, W., Sheffield, F. D., & Maccoby, N. Combining practice with demonstration in teaching complex sequences: serial learning of a geometric-construction task. In Student response on programmed instruction: symposium on programmed instruction. Chapter 4, National Academy of Sciences and National Research Council, Washington, D.C., 1961.

APPENDICES

APPENDIX A

PAIRED ASSOCIATE TASK: INSTRUCTIONS

The blocks are placed in their correct initial order (A, B, C, D, and 1, 2, 3, 4) in the holding tray (shown in Figure 1 and discussed in Chapter 4). Be sure that the "dot" blocks are in position in Row C and not in the matching slots (Row B). The pattern blocks are likewise in position in Row A in the order mentioned above from left to right. The order of presentation for subsequent trials is always calculated from the examiner's left to right. The box is closed and on the table in front of the examiner with the hinged side facing the subject. After the subject is seated across the table from the examiner, rapport is established and the basic information concerning the subject's name, date of birth, and age is completed at the top of each learning task form (Appendices B and D). The following instructions are given to introduce the tasks:

"We are going to play some learning games that will help us to understand how people learn things that they don't know. These won't be like games you have played before. . . . They will give you something brand new to learn. We just keep playing each new game until you have learned it and can do it twice without making any mistakes, and then we try another one. The games are in this box. Would you like to see the first one?"

Open the lid and lay it flat so that the tray and blocks are exposed. Since not all subjects have similar verbal abilities, concentrate as much as possible on the physical demonstration aspects of the explanation, touching and handling and sliding the blocks very surely and distinctly. Moving from left to right, slide each "dot" block from Row C into its correct slot (Row B) while saying:

"Every block in this top row goes together with this one in the bottom row, and they should always be put together just like I am doing now.

I want you to remember which one goes with which so you can always put the same two together when they are separated like this."

After all blocks are paired and located in Rows A and B, wait four seconds and then slowly slide the dot blocks back to the top of the tray (Row C). Then repeat a demonstration of only the first pair, A-1, saying:

"You have to be able to remember that this one goes with this one [slide dot Block A into position (Row B), touch Block 1, slide Block A back to the top of the tray (Row C)] so that when they are all mixed up like this [shuffle both rows of blocks into different order using both hands with the lid lowered] you will know that this one [slide Block A from a different location on Row C back into its correct slot above Block 1 in Row B] goes back together with this one. No matter how much I mix them up, you should always try to put the same sets of two blocks together."

While the subject is absorbing this, slowly slide dot blocks back into A, B, C, D position at the top of the tray (Row C).

"I am going to show you again which blocks belong together and then I will mix them all up so you can try to put them back together. Now watch carefully. . . ."

Beginning with A, slowly slide each block into its slot, touch its mate while saying:

"Now this one goes with this one. . . .
and THIS one goes with THIS one. . . ."

until all pairs are in position.

"Look carefully now so you will remember."

Time a pause of 10 seconds. Then say:

"Ready to try now?"

Raise lid to screen tray and arrange blocks for the first trial according to the arrangement at the top of column 1 on the paired associate score sheet. Lid is lowered. . . .

"See if you can remember which blocks go together."

If the subject seems confused or hesitant, encourage with:

"Try one as soon as you're ready. . . ."

Begin recording with the first response. Be thoroughly familiar with scoring instructions before beginning. When the child makes an error, stop him promptly by placing a hand over his hand and the block and pushing both slowly back into starting position, saying:

"No . . . that's not the one it belongs with. Try another one. . . ."

If the subject attempts to move the same block twice in succession, stop him again with your hand,

"When you don't get it right the first time, always try a different block the next time, ok?"

(Otherwise some children simply take one block at a time and "hopscotch" them over the slots until they find the match.)

Criterion for all subjects is reached with two consecutive errorless trials.

APPENDIX A

SEQUENTIAL LEARNING TASK: INSTRUCTIONS

Only Row B of the slotted black tray is used for this task. The "dot" and "pattern" blocks are to be removed completely from view, and the seven translucent white blocks in the kit are spread on the lid of the Black Box while the examiner explains:

> "Now we are going to play a tapping game. Which one of these blocks would you like to use to tap with?"

When the child chooses one, the examiner picks up another and arranges the remaining five in the slots on Row B (Figure 1).

> "I am going to hit these blocks. When I am done I want you to hit them just like I did. . . ."

Using a point of the block as the tapping surface, the examiner then taps out the first pattern (listed on the score sheet), using an even rhythm of ½ second per tap. The speed of presentation seems to make a difference, so it is important to practice even, correctly timed rhythm. Always tap from the subject's left to right when starting a pattern. Incorrect responses are followed by another demonstration of the pattern by the examiner with:

> "No, that's not quite right. Watch me again."

Promptly repeat the correct pattern for the subject to immediately re-try, and continue repeating in this fashion until the subject reproduces a correct sequence.

> "That's right, try it once more."

Demonstrate again immediately. Two consecutive, errorless trials are necessary for criteria. When a second correct sequence is achieved, say:

"That's right. You've learned that one very well. Now try a new one."

Continue to the next pattern and follow the same procedure with each of the remaining sequences. Responses are recorded during the task (dash for incorrect and a plus for correct responses) but do not hold the pencil in your hand with the block while demonstrating patterns.

APPENDIX B

EXAMINER'S SCORE SHEETS USED IN THE PRESENT STUDY

1. Paired Associate Task Score Sheet.

2. Sequential Learning Task Score Sheet.

(The combined Black Box Test learning curves were derived by combining the errors, trial by trial, across these two score sheets.)

PAIRED ASSOCIATE TASK SCORE SHEET

Name ______________________
Date of Birth ______________
Age ______________________
Date ______________________

Total No. of PA Errors ______
Composite PA Score
[200 - errors] ______

Composite PA Score	____
Composite SEQ Score	____
Total BBT Score	____
PA & SEQ Score/2	____

NUMBER OF RESPONSES

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1																						
2																						
3																						
4																						
5																						
6																						
7																						
8																						
9																						
10																						
11																						
12																						
13																						
14																						
15																						
16																						
17																						
18																						
19																						
20																						
21																						
22																						
# of Errors																						

TRIALS

SEQUENTIAL LEARNING TASK SCORE SHEET

Name ______________________

Sex ______________________

Examiner ______________________

Total Number of Sequential Errors ____

Composite Sequential Score [200 - errors] ______________________

																													Total Errors per Sequence
Total Errors per Trial																													

Total Sequential Errors

Comments:

APPENDIX C

MATERIALS UTILIZED FOR ADMINISTRATION OF THE MATH TASK

1. Math pre-test. (Administered to all children within the selected second-grade classroom.)

2. Study Guide. (The visual stimulus aligned with the tape presentation for each study session.)

3. Test Sheet A
4. Test Sheet B
5. Test Sheet C

(These are the answer sheets utilized by each student for each test session. Presentation orders were randomly determined.)

P R E - T E S T (M A T H)

NAME______________________________

1 1	5	1 0	8
X 6	X 6	X 6	X 6

____

1 2	6	2	7
X 6	X 6	X 6	X 6

____

4	1	9	3
X 6	X 6	X 6	X 6

S T U D Y   G U I D E

1	X	6	=	6
2	X	6	=	12
3	X	6	=	18
4	X	6	=	24
5	X	6	=	30
6	X	6	=	36
7	X	6	=	42
8	X	6	=	48
9	X	6	=	54
10	X	6	=	60
11	X	6	=	66
12	X	6	=	72

TEST SHEET A

NAME____________________ NO. OF ERRORS __________

DATE ____________________ SESSION NO. __________

---

1. 12 X 6 = ________
2. 3 X 6 = ________
3. 5 X 6 = ________
4. 6 X 6 = ________
5. 8 X 6 = ________
6. 4 X 6 = ________
7. 11 X 6 = ________
8. 2 X 6 = ________
9. 9 X 6 = ________
10. 1 X 6 = ________
11. 10 X 6 = ________
12. 7 X 6 = ________

TEST SHEET B

NAME ____________________ NO. OF ERRORS __________

DATE ____________________ SESSION NO. __________

1. 8 X 6 = ______

2. 12 X 6 = ______

3. 4 X 6 = ______

4. 6 X 6 = ______

5. 3 X 6 = ______

6. 1 X 6 = ______

7. 2 X 6 = ______

8. 10 X 6 = ______

9. 7 X 6 = ______

10. 5 X 6 = ______

11. 9 X 6 = ______

12. 11 X 6 = ______

TEST SHEET C

NAME______________________ NO. OF ERRORS __________

DATE ______________________ SESSION NO. __________

1.	3	X	6	=	________
2.	2	X	6	=	________
3.	5	X	6	=	________
4.	4	X	6	=	________
5.	7	X	6	=	________
6.	1	X	6	=	________
7.	10	X	6	=	________
8.	6	X	6	=	________
9.	11	X	6	=	________
10.	12	X	6	=	________
11.	9	X	6	=	________
12.	8	X	6	=	________

APPENDIX D

MATERIALS UTILIZED FOR ADMINISTRATION OF THE WRITING TASK

1. Writing pre-test.	(Administered to each student within the selected second-grade classroom.)
2. Cursive letter criteria.	(Detailed explanation of the criteria established to determine the accuracy of each of the eight selected letters.)

WRITING TASK PRE-TEST

a	b	c	d	e	f	g	h
i	j	k	l	m	n	o	p
q	r	s	t	u	v	w	x
y	z						

CURSIVE LETTER CRITERIA

All letters were corrected according to the following criteria:

1. Continuous. All letters must be made in one motion. It was considered an error for the subject to take the pencil off the paper while completing the letter.

2. Overwork. Unnecessary or inappropriate retracing (allowed stem [2] of letter d) is considered an error.

3. Erasures. Any erasures were automatically considered an errored letter.

4. Time. All letters must be completed within seven second.

5. Slant. Any letters slanting towards the left were considered as errors. Appropriate slanting ranged from perpendicular to 30 degrees to the right. This was measured on a protractor (90-120 degrees) using the longest straight line of each letter as the base indicator.

6. X Indicator. Greater tha n 1/32" at any X indicator (distance between the letter and the guide line) for each letter was an error.

7. Irregularities. Major irregularities or angle substitution for curves were considered an error.

8. Direction. All letters must be made according to the direction demonstrated (attention to the starting and finishing points of each letter).

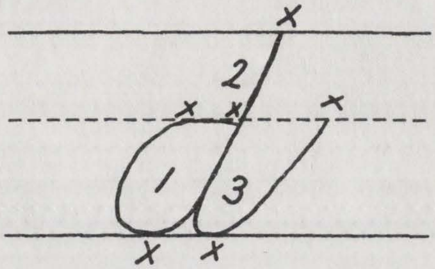
The following specific criteria monitored the accuracy of each individual letter. (Note: The following diagrams are enlarged to provide additional clarity.)

<u>CRITERIA DESCRIPTIONS</u> — <u>LETTER DIAGRAMS</u>

1. Width (1) not less than 1/8" nor greater than ¼".

2. No separation of 2 (stem).

3. Width (3) not less than 1/16" nor greater than 5/16" at largest point.



---

1. Width (1) not less than 1/8" nor greater than ¼".

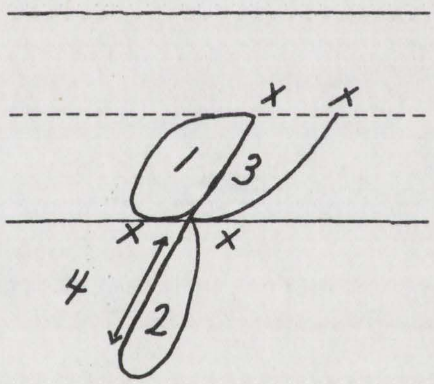
2. Width (2) not less than 1/16" nor greater than 3/16" at greatest point.

3. Width (3) not less than 1/16" nor greater than 5/16" at largest point.

4. Length (4) not less than 3/16" nor greater than 7/16".

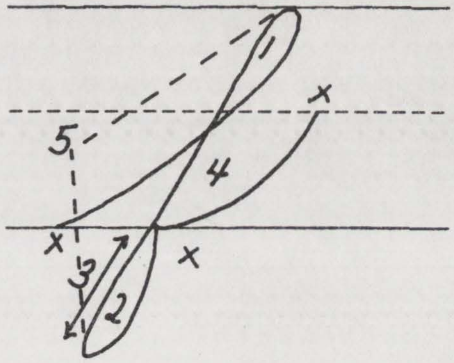
5. Width (2) not larger than width (1).

6. Loop (1) must touch at least two midline dots before descending to bottom guideline.



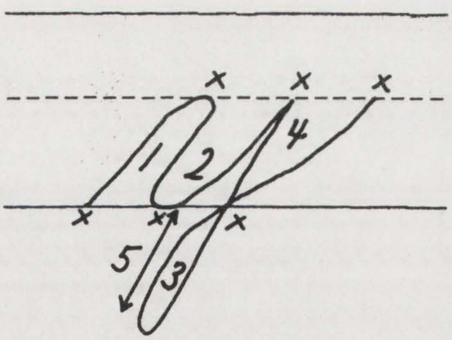
CRITERIA DESCRIPTIONS	LETTER DIAGRAMS

1. Width (1) not less than 1/16" nor greater than 3/16" at largest point.
2. Width (2) not less than 1/16" nor greater than 3/16" at largest point.
3. Length (3) not less than 3/16" nor greater than 7/16".
4. Width (4) not less than 1/16" nor greater than 5/16" at greatest point.
5. Line (5) straight, no major irregularities or changes in direction.
6. Loop (1) not less than 1/16" from the top guideline.
7. Not greater than 1/16" difference between width (1) and (2).



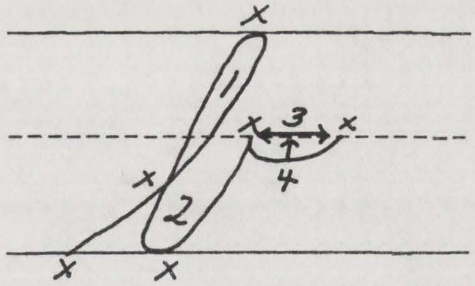
---

1. Width (1) not less than 1/16" nor greater than 5/16" at largest point.
2. Width (2) not less than 1/8" nor greater than 5/16" at largest point.
3. Width (3) not less than 1/16" nor greater than 3/16" at largest point.
4. Width (4) not less than 1/16" nor greater than 5/16" at largest point.
5. Length (5) not less than 3/16" nor greater than 7/16" at largest point.



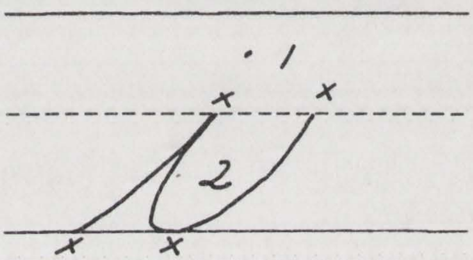
CRITERIA DESCRIPTIONS | LETTER DIAGRAMS

1. Width (1) not less than 1/16" nor greater than 3/16" at largest point.
2. Width (2) not less than 1/32" nor greater than 3/16" at any point.
3. Width (3) not less than 1/32" nor greater than 1/4".
4. Depth (4) not less than 1/32" nor deeper than 1/8".



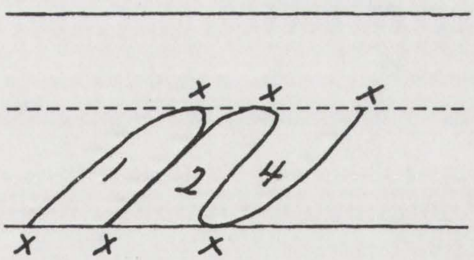
---

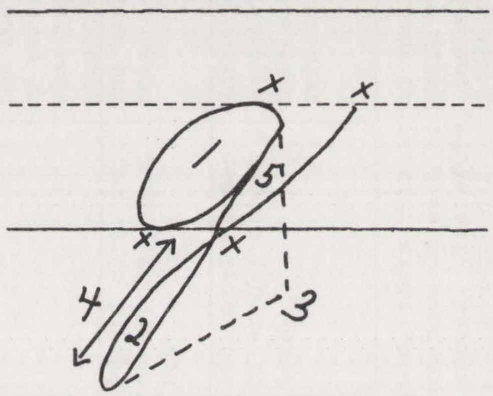
1. (1) must occur between the top guideline and middle line.
2. Width (2) not less than 1/16" nor greater than 5/16" at largest point.



---

1. Width (1) not less than 1/16" nor greater than 1/4" at largest point.
2. Width (2) not less than 1/16" nor greater than 1/4" at largest point.
3. Not greater than 1/16" difference between width (1) and (2).
4. Width (4) not less than 1/16" nor greater than 5/16" at largest point.
5. Midline must join greater than or equal to 1/8".



CRITERIA DESCRIPTIONS	LETTER DIAGRAMS
1. Width (1) not less than 1/8" nor greater than 1/4" at largest point.	
2. Width (2) not less than 1/16" nor greater than 3/16" at largest point.	
3. Line (3) straight, no major irregularities or changes in direction.	
4. Length (4) not less than 3/16" nor greater than 7/16".	
5. Width (5) not less than 1/16" nor greater than 5/16" at largest point.	
6. Width (2) not greater than width (1).	
7. Loop (1) must touch at least 2 midline dots before descending to bottom guideline.	

APPENDIX E

SUMMARY OF TASKS

Task	Trial Definition	Task Criterion	Calculation of Composite Scores	Data Collection System for Learning Curves Presented in This Study	Ceiling Levels
Paired Associate	Includes the No. of R's necessary to correctly associate 4 sets of "dot"-"pattern" stimuli.	Two consecutive errorless trials (x number of errors/trial).	200-total number of errors for task.	Not demonstrated.	44 trials
Sequential*	Each attempt to reproduce a tapping pattern.	Two consecutive errorless trials (for each of 5 tapping sequences). One trial = 1 error.	200-total number of errors across task.	Not demonstrated.	30 trials for each tapping pattern.
Combined BBT Performance	Trial by trial combination of PA & Seq. tasks. (e.g., BBT Trial 1 = sum of errors on Trial 1 of PA and Trial 1 of Sequential.)	Fulfillment of PA & Seq. task criteria.	Sum of PA & Seq. composite scores divided by 2.	Number of errors per trial by combining respective trials of PA and Seq. tasks to criteria. (The errors per trial for the Seq. task = the sum of errors across the respective trials on all 5 tapping sequences.	Subject to PA & Seq. levels.
Math	Each study-guide-test session.	Two consecutive errorless trials (1-12 errors per trial).	400-total number of errors across task.	The number of errors per trial to task criterion.	60 trials
Writing*	Each attempt to reproduce a letter.	Two consecutive errorless (for each of 8 selected letters. One trial = one error.	500-total number of errors across task.	Number of errors per trial by by combining respective trials across all 8 letters to task criterion.	60 trials per letter.

*Both the writing and the sequential learning tasks consist of smaller sub-tasks (eight letters and five tapping sequences, respectively). Performances on each of these sub-tasks were combined, trial by trial, to determine their respective learning curves. Therefore, the eight letters were considered one task as were the five tapping sequences.

APPENDIX F

DATA PROVIDING COMPOSITE SCORE ANALYSIS FOR TABLE 1 AND FIGURES 3 AND 4

SUBJECTS	PAIRED ASSOCIATE			SEQUENTIAL			COMBINED BLACK BOX TEST				MATH (MULTIPLICATION FACTS)					WRITING (EIGHT LETTERS)					TEACHER RATING (RANK)
	Errors	Comp. Score	Rank	Errors	Comp. Score	Rank	Comp. Score	Rank	Z Score	T Score	Errors	Comp. Score	Rank	Z Score	T Score	Errors	Comp. Score	Rank	Z Score	T Score	
S1	21	179	6	11	189	3.5	184	6	51.6	51	138	262	6	52.2	51	82	418	8	52.7	46.5	4
S2	46	154	8	10	190	2	172	8	52.0	46.5	123	277	5	54	53	99	401	9	51.1	44	9
S3	15	185	5	11	189	3.5	187	3.5	57.2	56.5	109	291	4	55.5	55.5	47	453	2	56	61.5	3
S4	107	93	11	39	161	11	127	11	36.5	38.5	199	201	9	45.4	44	128	372	10	48.4	41.5	8
S5	7	193	3	19	181	7	187	3.5	57.2	56.5	140	260	7	52	48.5	74	426	6	53.5	51	11
S6	77	123	10	20	180	8	151.5	9	44.9	44	281	119	11	36.3	38.5	65	435	4	54.3	55.5	5
S7	148	52	12	36	164	10	108	12	29.9	32	146	254	8	51.3	46.5	419	81	12	21	32	10
S8	1	199	2	26	174	9	186.5	5	57.0	53	63	337	3	60.6	58	73	427	5	53.6	53	2
S9	37	163	7	17	183	6	173	7	52.3	48.5	309	91	12	33.1	32	77	423	7	53.2	48.5	12
S10	0	200	1	13	187	5	193.5	1	59.4	67.5	58	342	2	61.2	61.5	56	444	3	55.2	58	6
S11	13	187	4	4	196	1	191.5	2	58.7	61.5	52	348	1	61.8	67.5	22	478	1	58.4	67.5	1
S12	73	127	9	60	140	12	133.5	10	38.7	41.5	276	124	10	36.8	41.5	191	309	11	42.5	38.5	7

APPENDIX G

DATA TABLES CONTAINING SUBJECT PERFORMANCE ACROSS TASKS

1. Subjects' raw scores. (Number of errors for each subject across the trials of each task.)

2. Mean and standard deviation data points. (Presented for each task.)

3. Percentage distribution. (Percentage of trials ranging within the indicated standard deviations, for each subject and task.

4. Standard scores. (Based upon the length of the learning curves for each task. Single subject comparisons were provided by the relationship of these scores.)

APPENDIX G

INDIVIDUAL SUBJECT RAW SCORES (NUMBER OF ERRORS) ACROSS ALL TRIALS FOR EACH TASK

TASKS	SUBJECTS	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
BLACK BOX TEST	S1	4	5	3	2	2	2	1	3	6	2	1	1	0	0																																														
	S2	6	6	6	3	0	2	1	2	6	5	0	2	4	1	1	3	2	1	1	4	0	0																																						
	S3	5	13	2	2	2	2	1	0	0																																																			
	S4	7	4	6	7	4	15	8	11	4	10	4	8	2	4	3	2	6	3	8	2	3	1	2	4	2	2	1	5	1	0	4	2	1	0	0																									
	S5	3	3	2	4	3	2	2	3	2	2	0	0																																																
	S6	4	3	4	5	3	6	4	11	4	3	23	2	4	8	0	2	1	3	0	1	2	1	2	1	0	0																																		
	S7	9	3	5	4	5	6	7	8	4	3	9	3	3	3	13	4	4	6	2	4	5	2	8	5	5	4	2	2	1	7	3	2	8	7	1	2	1	2	2	1	7	2	0	0																
	S8	4	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	0																																					
	S9	3	4	6	5	3	2	3	3	2	2	2	3	5	2	1	2	2	1	0	1	1	1	0	0																																				
	S10	2	2	2	1	1	1	1	1	1	1	0	0																																																
	S11	2	5	2	2	1	1	1	2	1	0	0																																																	
	S12	8	5	11	8	7	4	3	6	6	3	2	6	5	5	10	10	5	2	3	3	4	2	2	2	2	2	2	2	2	2																														
MATH TASK	S1	12	12	8	9	9	9	9	9	9	6	6	7	5	6	3	2	5	3	2	2	3	0	2	0	0																																			
	S2	11	11	9	9	8	8	6	8	6	5	6	4	5	6	5	4	2	2	3	1	2	0	1	0	1	0	0																																	
	S3	11	10	12	9	9	6	7	6	7	7	4	3	6	4	3	3	2	0	0																																									
	S4	12	12	12	11	12	12	12	10	9	7	9	8	9	9	7	6	6	4	4	4	3	3	5	2	3	3	3	0	2	0	0																													
	S5	11	10	11	10	11	9	9	9	8	7	7	7	6	4	4	5	3	2	3	0	1	2	0	0																																				
	S6	10	10	11	12	10	10	9	8	7	8	10	8	10	9	9	10	9	7	7	8	8	8	7	5	7	5	7	7	5	4	5	4	4	5	3	4	3	3	3	2	2	0	0																	
	S7	12	12	10	9	8	8	8	7	8	8	7	5	6	5	4	5	5	3	4	4	1	3	2	1	1	0	0																																	
	S8	10	5	8	6	9	7	5	0	5	1	4	2	0	1	0	0																																												
	S9	12	12	12	10	11	11	11	11	9	10	9	10	10	11	8	7	8	7	6	5	8	9	7	7	8	8	7	6	5	6	5	4	4	4	5	4	3	2	3	3	2	2	1	0	3	2	0	1	0	0										
	S10	8	8	8	7	3	5	5	4	3	1	3	3	0	0																																														
	S11	10	11	9	8	2	7	3	0	2	0	0																																																	
	S12	11	11	9	9	9	10	10	9	8	10	8	8	9	9	8	8	7	7	9	6	6	6	4	6	4	5	4	6	6	6	5	6	3	3	2	4	4	3	4	2	3	3	3	1	2	0	0													
WRITING TASK	S1	8	7	6	5	5	4	4	4	3	4	4	4	3	4	3	1	0	1	1	1	1	1	1	0	1	1	0	1	1	1	0	1	0	1	0	0																								
	S2	8	7	4	6	4	5	5	5	3	4	5	4	4	5	3	3	4	3	3	2	3	2	2	2	1	1	1	0	0																															
	S3	8	5	6	6	6	6	4	2	2	1	1	0	0																																															
	S4	8	6	8	6	7	8	6	7	8	8	5	6	5	5	4	4	3	2	3	1	1	2	2	2	2	2	1	2	1	1	1	1	0	0																										
	S5	7	7	6	5	4	3	5	5	3	5	3	4	5	2	2	1	1	1	1	1	1	1	0	1	0	0																																		
	S6	7	4	7	7	4	4	3	3	4	3	3	3	2	2	3	2	0	1	1	1	0	1	0	0																																				
	S7	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	7	7	8	8	7	8	7	7	7	7	7	6	7	7	7	7	7	6	7	6	7	7	7	7	7	7	7	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	
	S8	7	6	6	6	4	4	3	3	3	2	2	1	2	2	2	2	2	1	2	1	2	2	1	1	1	0	1	1	1	1	0	1	0	0																										
	S9	8	6	7	7	5	5	6	5	4	4	4	3	1	1	1	1	1	2	2	1	2	0	1	0	0																																			
	S10	7	7	4	5	5	3	4	2	4	3	4	3	2	2	1	0	0																																											
	S11	8	6	3	3	1	1	0	0																																																				
	S12	8	8	8	7	7	6	7	7	6	6	6	7	7	6	7	5	4	5	5	5	5	4	4	5	4	4	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

TRIALS

[Ceiling Level, Seq. Task: 2 patterns]

[Ceiling Level: 6 Letters] ← (Writing Task, S7)

[Ceiling Level: 1 Letter] ← (Writing Task, S12)

APPENDIX G

CALCULATED DATA POINTS FOR THE MEAN AND STANDARD DEVIATION CURVES ACROSS EACH TASK

TASKS	Standard Deviation and Mean Curves	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
BBT	1 Sd > $\overline{X}$	7.1	7.5	7.0	6.0	4.7	7.7	5.3	8.1	5.3	5.5	11	5.0	4.1	4.5	5.5	4.9	3.9	3.2	3.6	2.9	3.1	1.4	3.5	2.5	2.3	2.0	1.2	2.3	1	2.8
	$\overline{X}$	4.8	4.6	4.3	3.8	2.7	3.7	2.8	4.3	3.1	2.7	4.3	2.2	2.0	2.0	2.4	2.0	1.8	1.4	1.2	1.3	1.3	.6	1.2	1.0	.8	.7	.4	.8	.3	.8
	1 Sd < $\overline{X}$	2.5	1.7	1.6	1.6	.7	0	.3	.5	.9	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
MATH	1 Sd > $\overline{X}$	12	12.4	12.9	10.7	11.4	10.6	10.5	10.4	9.1	9.2	9	8.4	8.5	9.1	7.5	7.5	7.1	5.7	6.2	5.3	5.8	5.9	5.0	4.5	4.9	4.6	4.6	4.5	3.9	3.8
	$\overline{X}$	10.8	10.3	9.9	9.1	8.4	8.5	7.8	6.8	6.8	5.8	6.1	5.4	5.5	5.3	4.3	4.2	3.9	2.9	3.2	2.5	2.7	2.6	2.3	1.8	2.0	1.8	1.8	1.6	1.5	1.3
	1 Sd < $\overline{X}$	9.6	8.2	6.9	7.5	5.4	6.4	5.1	3.2	4.5	2.4	3.2	2.4	2.5	1.5	1.1	.9	.7	.1	.2	–	–	–	–	–	–	–	–	–	–	–
WRITING	1 Sd > $\overline{X}$	8.2	7.6	7.8	7.2	6.9	6.8	6.7	7.1	6.3	6.5	6.0	6.1	5.9	5.6	5.3	4.7	4.1	4.0	4.6	4.2	4.0	4.1	3.6	3.8	3.4	3.5	3	2.8	3	3
	$\overline{X}$	7.7	6.4	6.1	5.9	5.0	4.8	4.6	4.3	4.0	4.0	3.8	3.6	3.3	3.1	2.8	2.3	1.8	1.9	2.2	1.8	1.8	1.8	1.5	1.5	1.3	1.3	1	1	1	1
	1 Sd < $\overline{X}$	7.2	5.2	4.4	4.6	3.1	2.8	2.5	1.5	1.7	1.5	1.6	1.1	.7	.6	.3	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–

TASKS	Standard Deviation and Mean Curves	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
BBT	1 Sd > $\overline{X}$	.9	1.1	3.1	2.6	.4	.8	.4	.8	.8	.4	2.6	.8	0	0																
	$\overline{X}$	.6	.3	.8	.6	.1	.2	.1	.2	.2	.1	.6	.2	0	0																
	1 Sd < $\overline{X}$	–	–	–	–	–	–	–	–	–	–	–	–	–	–																
MATH	1 Sd > $\overline{X}$	3.6	3.4	2.6	2.9	2.4	2.6	2.3	1.9	2.3	1.7	1.7	1.4	1.2	.8	1.4	.8	0	.4	0	0										
	$\overline{X}$	1.3	1.2	.9	1	.83	1	.83	.67	.83	.58	.58	.42	.33	.08	.42	.17	0	.08	0	0										
	1 Sd < $\overline{X}$	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–		–	–	–										
WRITING	1 Sd > $\overline{X}$	2.8	2.9	2.7	2.4	2.7	2.3	2.7	2.7	2.7	2.7	2.7	2.7	2.7	2.3	2.3	2.3	2.3	2.3	2.3	2.3	2.3	2.3	2.3	2.3	2.3	2.3	2.3	2.3	2.3	2.3
	$\overline{X}$	.83	.92	.7	.7	.7	.6	.7	.7	.7	.7	.7	.7	.7	.6	.6	.6	.6	.6	.6	.6	.6	.6	.6	.6	.6	.6	.6	.6	.6	.6
	1 Sd < $\overline{X}$	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–

APPENDIX G

SUMMARY TABLE OF THE PERCENTAGE DISTRIBUTION ACROSS INDICATED STANDARD DEVIATIONS FOR EACH SUBJECT AND TASK

Standard Deviation	Task	SUBJECTS											
		S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12
$\geq$ 1 Sd above $\bar{X}$	BBT	8	9	11	40	0	12	63	0	4	0	0	57
	Math	4	0	0	25	0	70	4	0	74	0	0	60
	Writing	0	0	0	15	0	0	98	0	0	0	0	30
Within 1 Sd > $\bar{X}$	BBT	7	32	11	40	17	42	30	0	30	0	9	40
	Math	48	19	32	65	54	20	52	6	18	0	9	32
	Writing	22	72	31	62	23	13	2	15	36	12	12	70
Within 1 Sd < $\bar{X}$	BBT	85	55	56	14	83	38	2	100	68	83	82	3
	Math	48	81	52	10	46	10	44	31	2	21	36	6
	Writing	64	24	31	15	73	75	0	70	52	64	12	0
$\leq$ 1 Sd below $\bar{X}$	BBT	0	4	22	0	0	0	0	0	0	17	9	0
	Math	0	0	16	0	0	0	0	63	0	79	55	0
	Writing	0	4	38	0	4	8	0	6	0	12	76	0
On $\bar{X}$	BBT	0	0	0	6	0	8	5	0	8	0	0	0
	Math	0	0	0	0	0	0	0	0	6	0	0	2
	Writing	14	0	0	8	0	4	0	9	12	12	0	0

APPENDIX G

STANDARD SCORES BASED UPON THE LENGTH OF THE LEARNING FOR ALL TASKS

Subjects	Black Box Test	Math	Score	Writing	Score
S1	42.6	47.7	+	53.8	-
S2	50	49.2	+	50.6	+
S3	38	43	+	39.4	+
S4	62	52.3	+	52.5	+
S5	40.7	46.9	+	47.5	+
S6	53.8	61.7	+	46.2	-
S7	70.4	49.2	-	68.8 (Ceiling Level)	+
S8	50.9	40.6	_	52.5	+
S9	51.9	67.2	-	46.9	-
S10	40.7	39.1	+	41.9	+
S11	39.8	36.7	+	36.2	+
S12	57.4 (Ceiling Level)	64.8	+	68.8 (Ceiling Level)	+

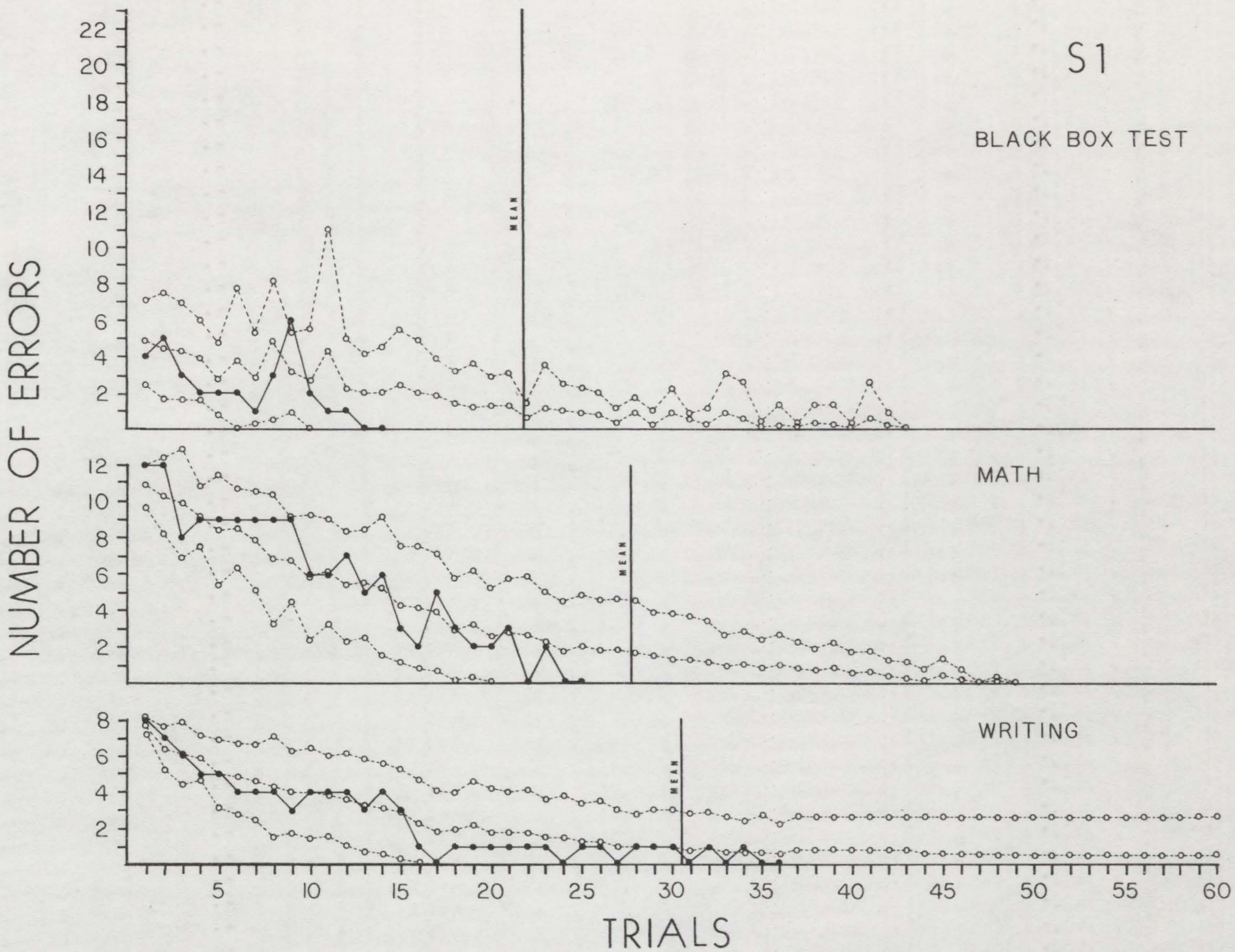
+ indicates a positive relationship between the Black Box Test curve length and task.

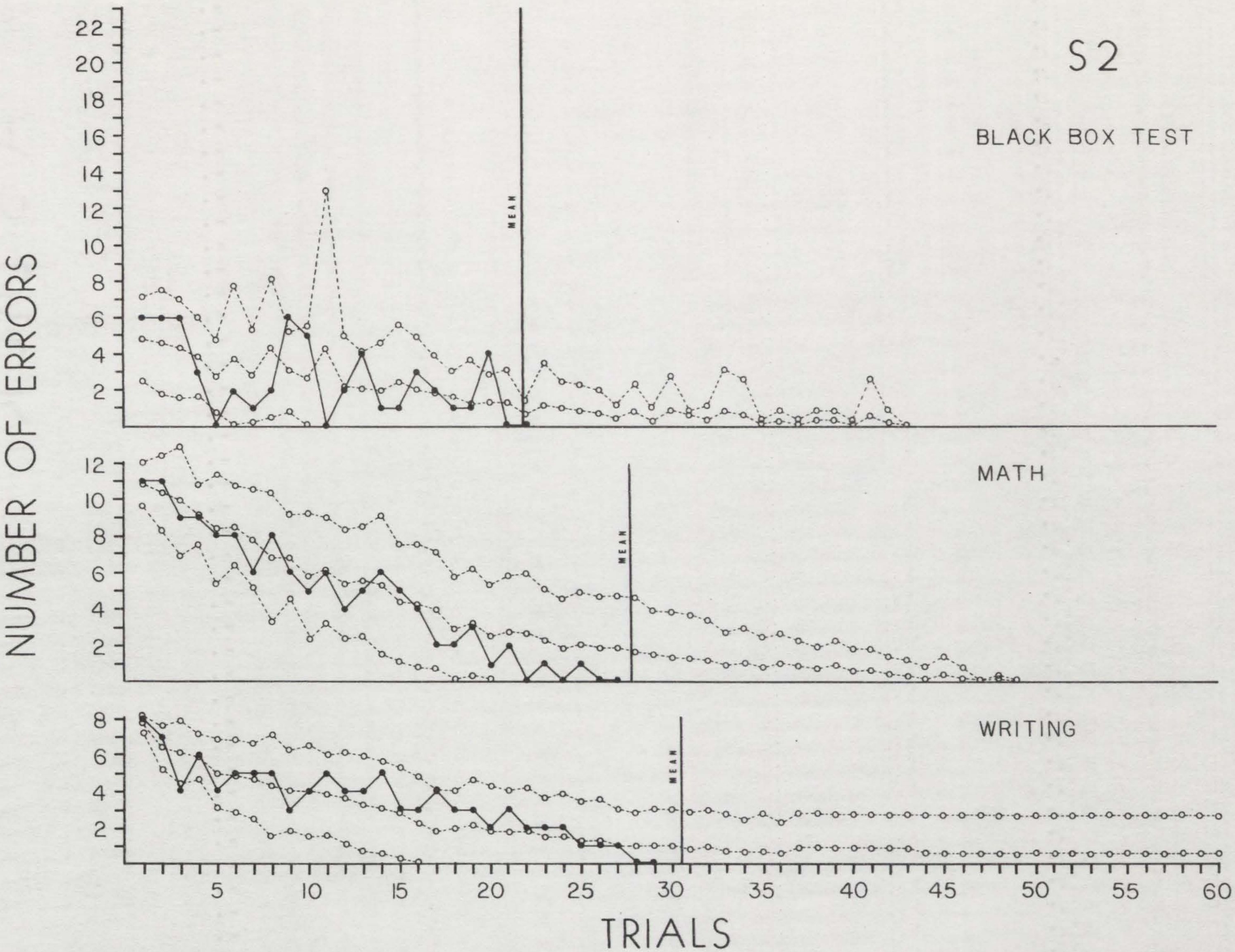
- indicates lack of a positive relationship between the Black Box Test curve length and task.

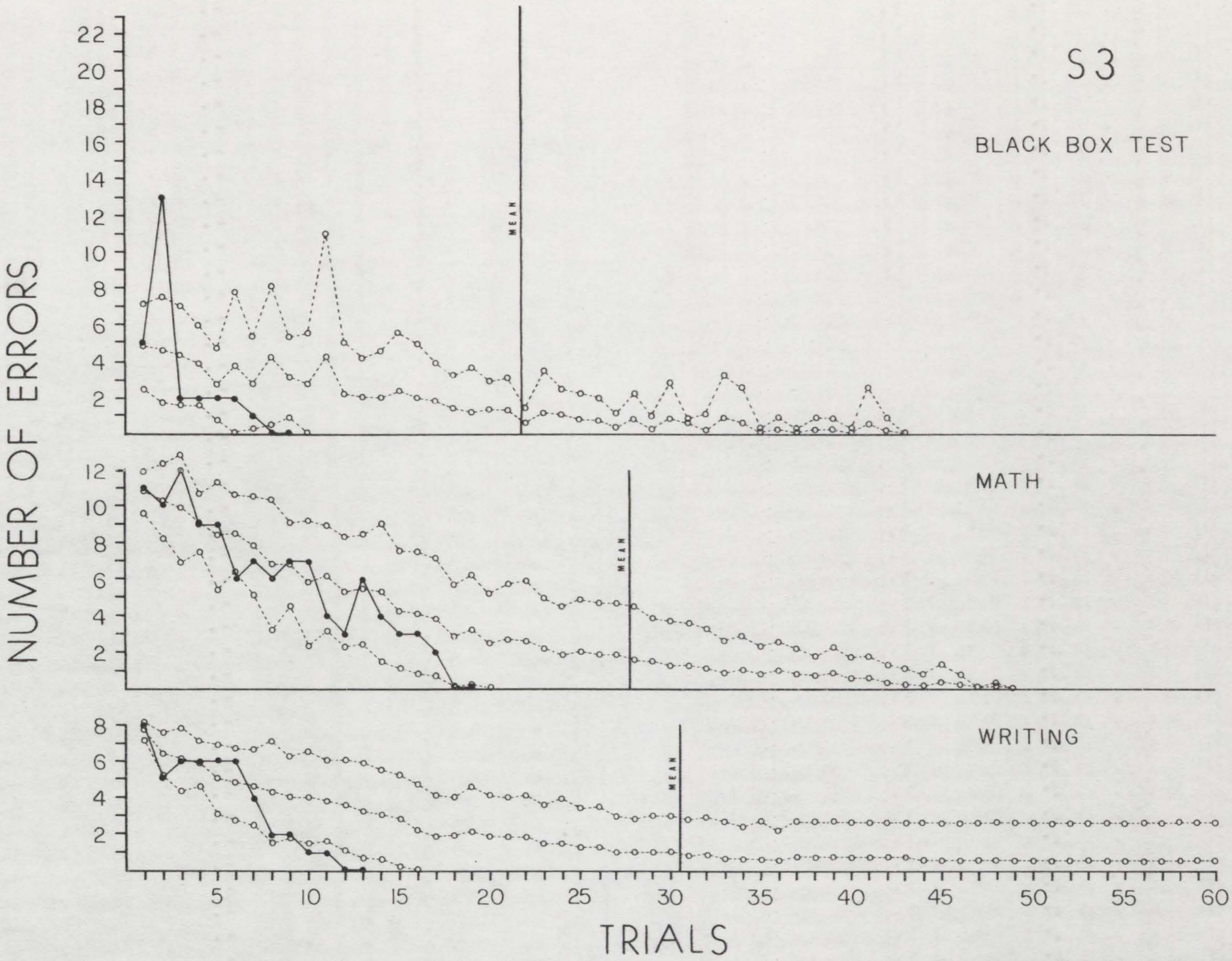
APPENDIX H

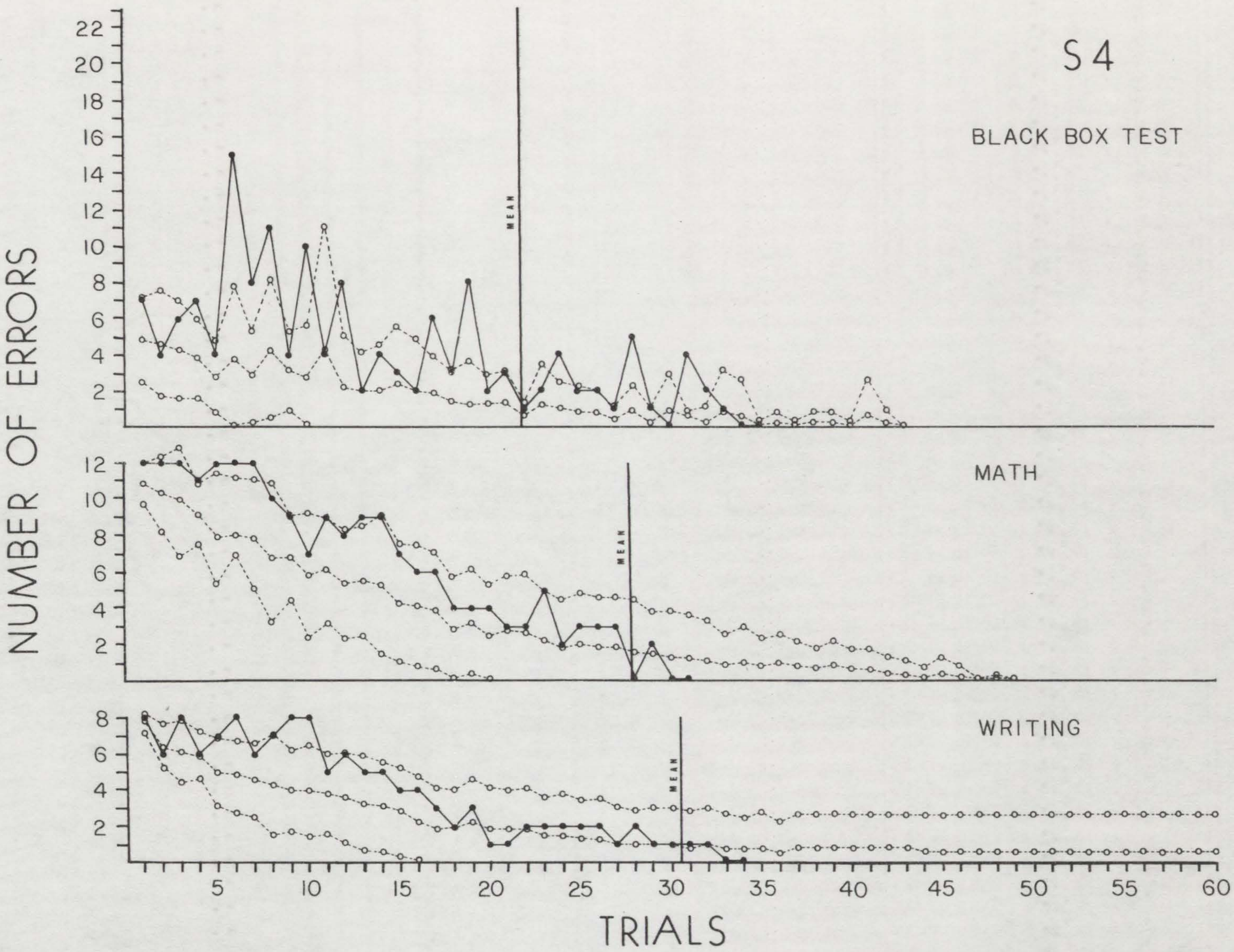
INDIVIDUAL LEARNING CURVES

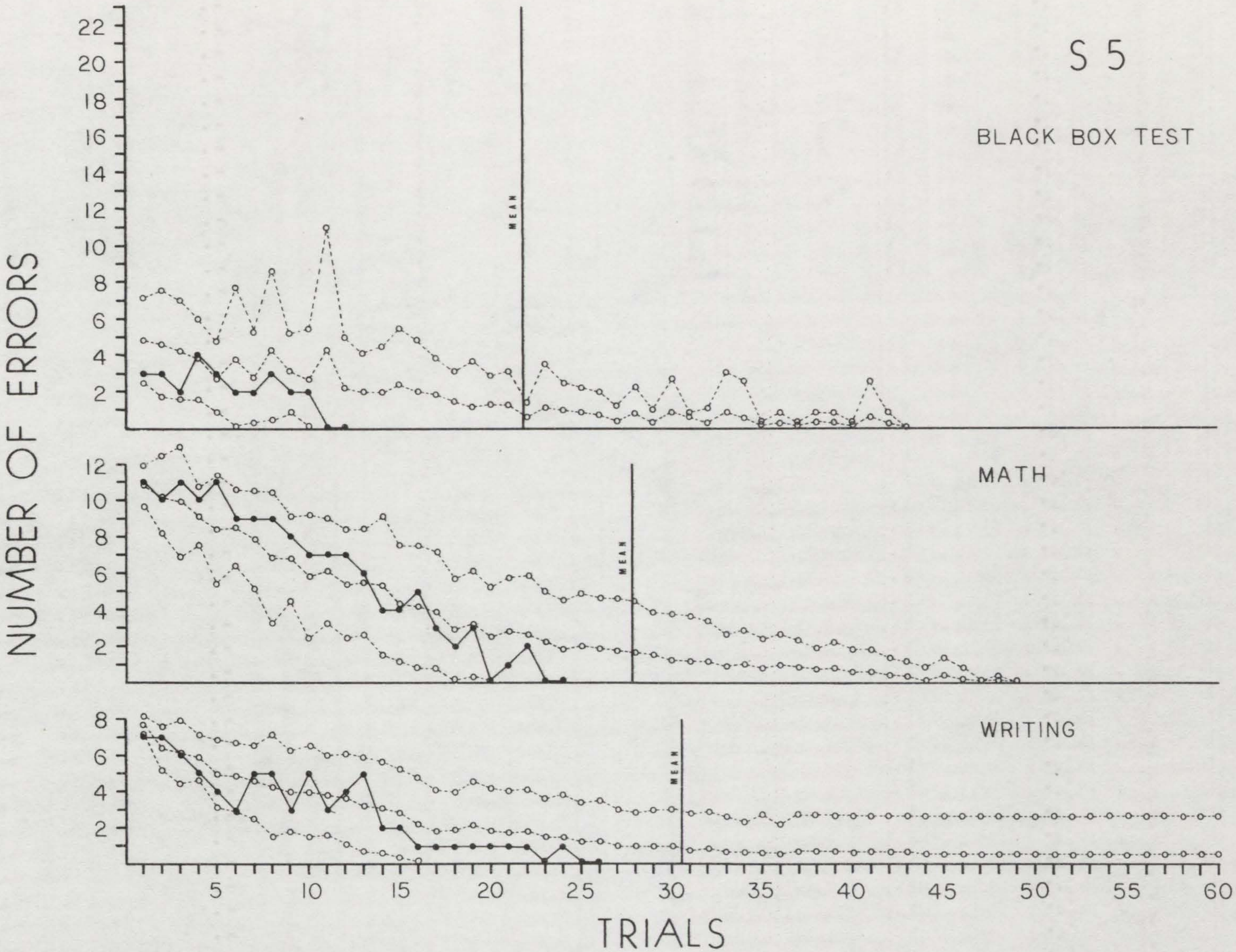
The following graphs, individually designated as Subjects 1 through 12, present the individual learning curves for all subjects and tasks. Each data point represents the number of errors (labelled along the ordinate) per trial (indicated on the abscissa) for each task (labelled at the top right-hand corner of each graph). The subject number is indicated at the top right-hand corner of each page. The average number of trials for each task is indicated by a vertical line connecting the abscissa. The subjects' individual learning curves are represented by a solid line connecting solid dots for each task. The mean and standard deviations were individually calculated on each trial for all tasks and are represented by dotted lines connecting circles. The middle dotted line represents one standard deviation above and below the mean, respectively. Those subject responses falling above the mean indicate trials containing more than the average number of errors for that subject and task. The opposite is true for those responses falling below the mean learning curve.


S1
BLACK BOX TEST
MATH
WRITING
MEAN
NUMBER OF ERRORS
TRIALS
22
20
18
16
14
12
10
8
6
4
2
5
10
15
20
25
30
35
40
45
50
55
60

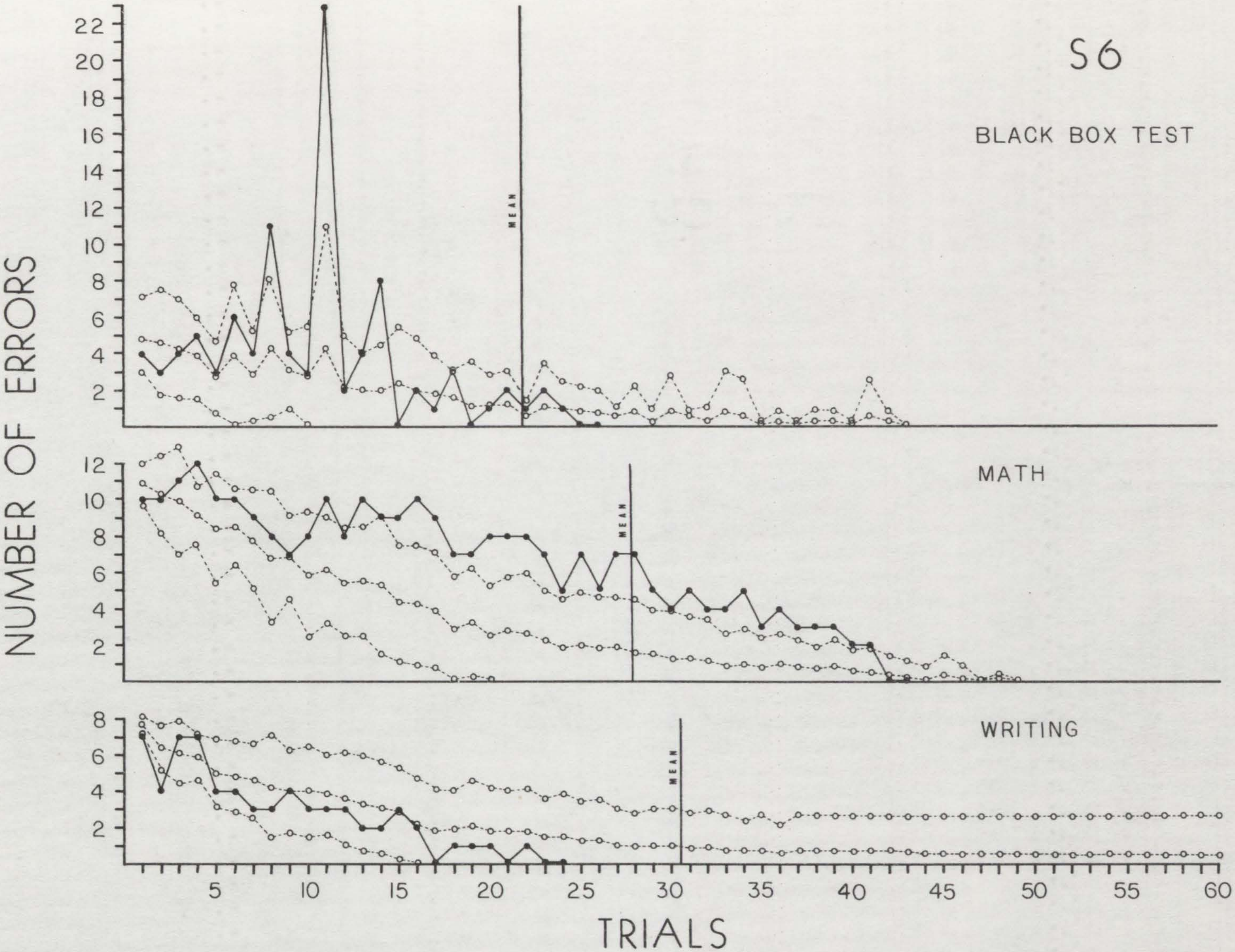

S2
BLACK BOX TEST
MATH
WRITING
MEAN
MEAN
MEAN
NUMBER OF ERRORS
TRIALS

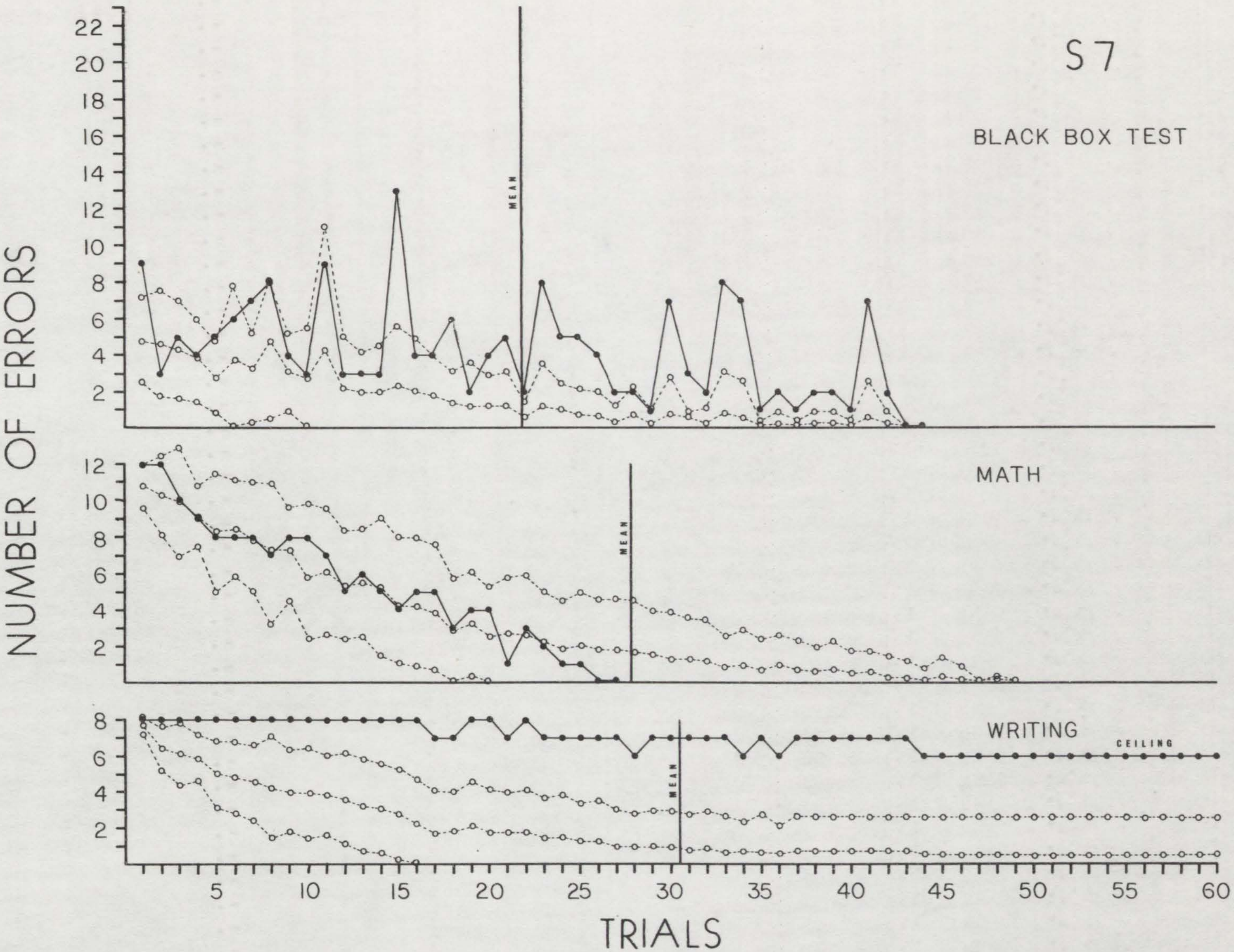

S3
BLACK BOX TEST
MATH
WRITING
MEAN
MEAN
MEAN
NUMBER OF ERRORS
22
20
18
16
14
12
10
8
6
4
2
12
10
8
6
4
2
8
6
4
2
5
10
15
20
25
30
35
40
45
50
55
60
TRIALS

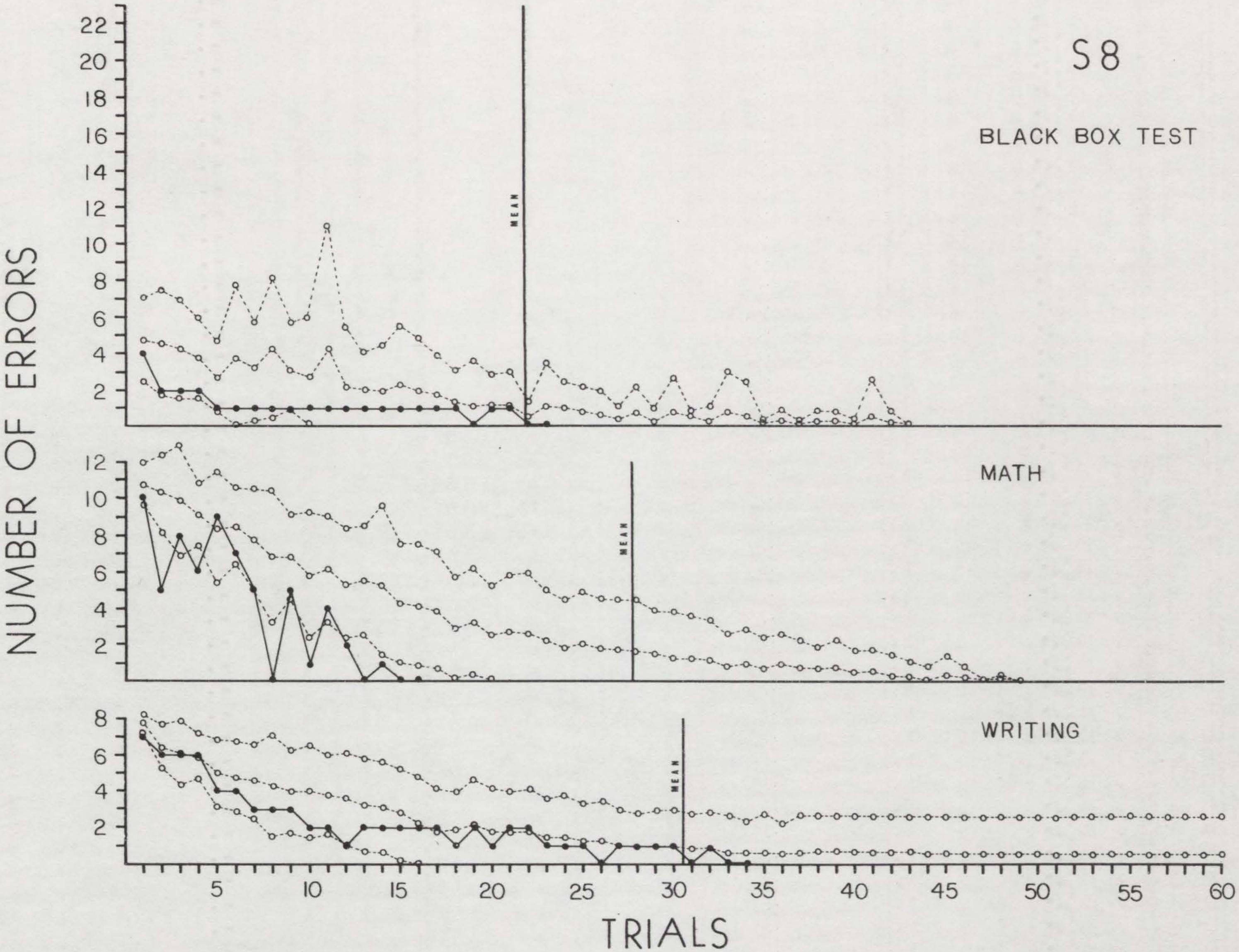

S4
BLACK BOX TEST
MATH
WRITING
MEAN
NUMBER OF ERRORS
TRIALS
22
20
18
16
14
12
10
8
6
4
2
5
10
15
20
25
30
35
40
45
50
55
60

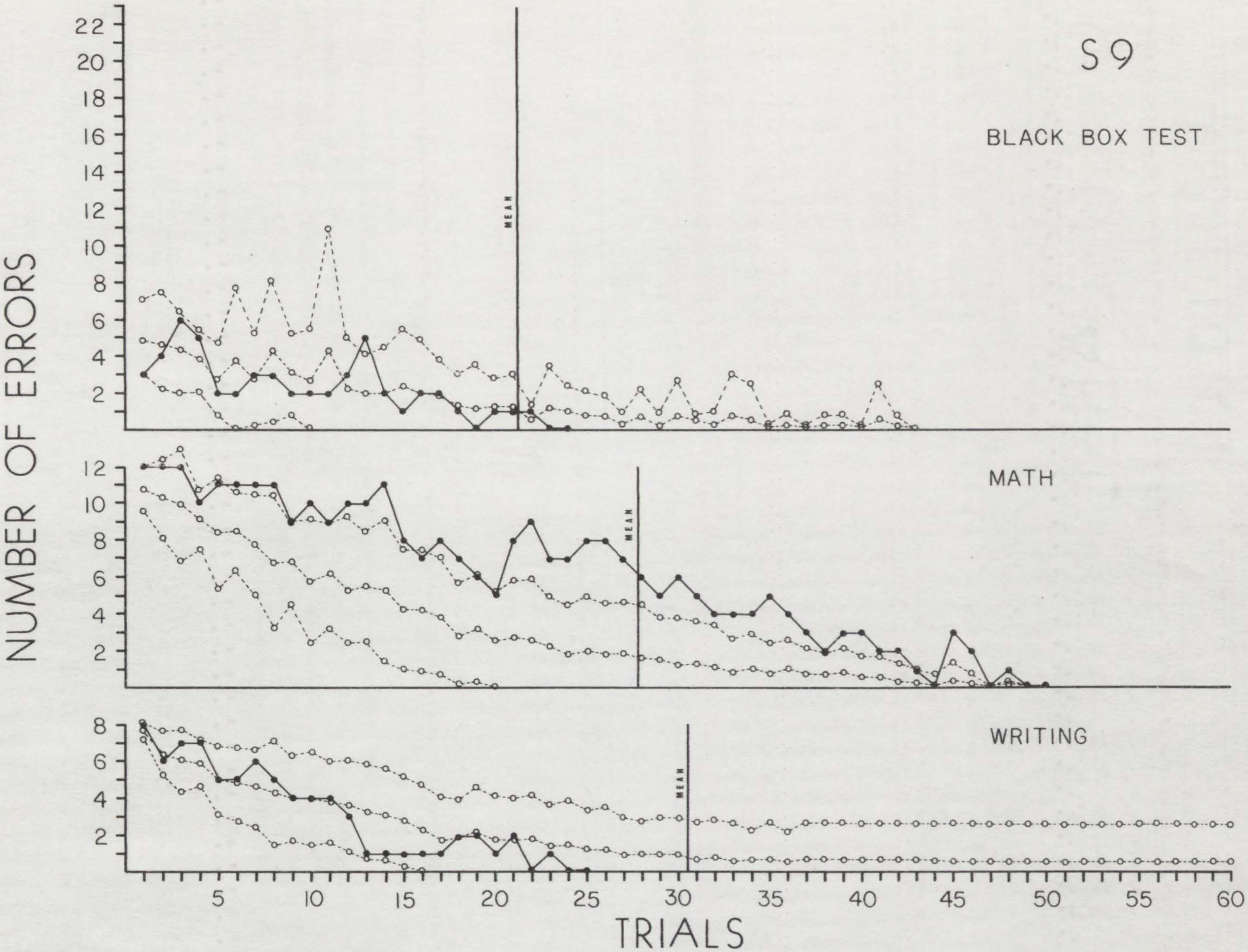


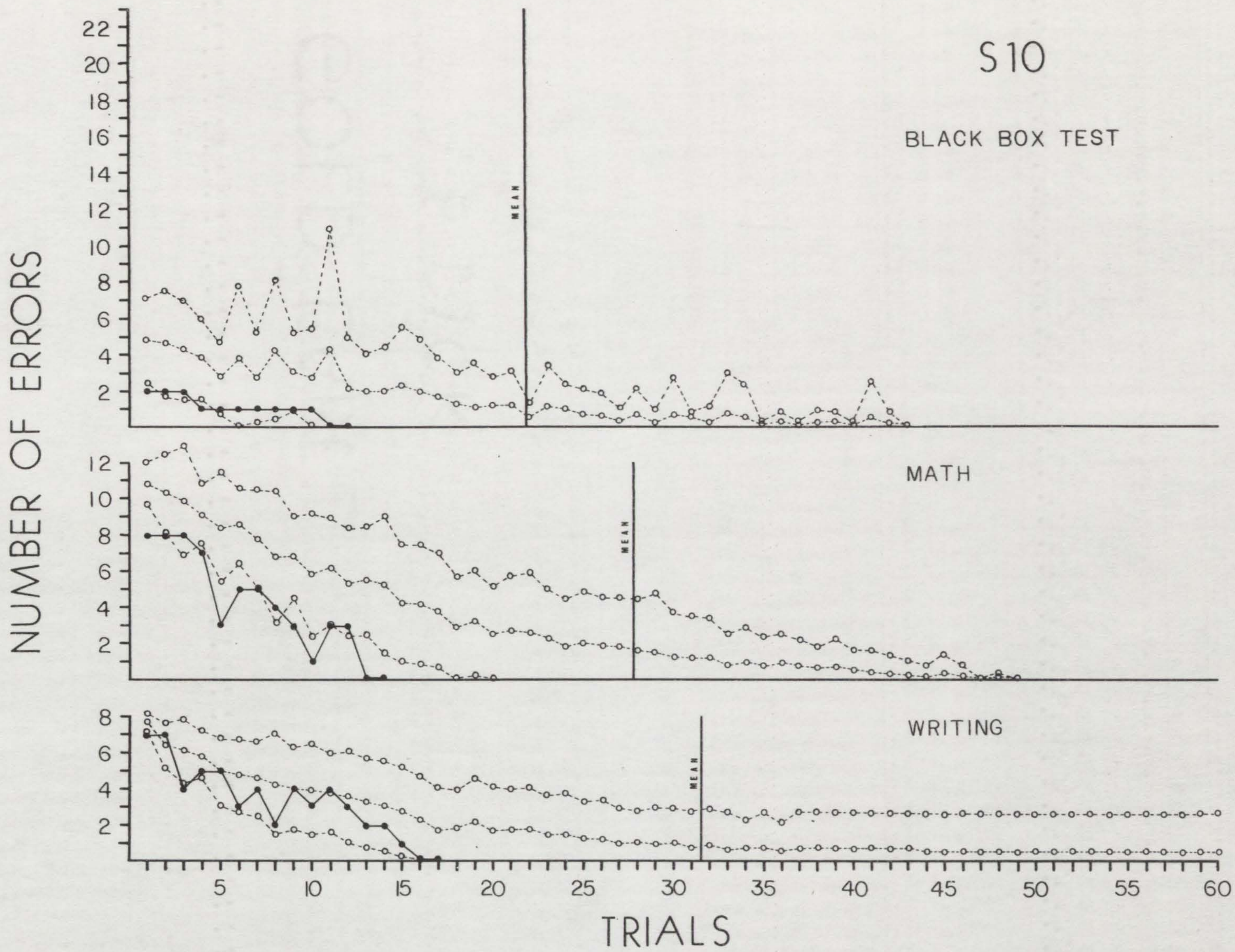
S 5
BLACK BOX TEST
MEAN
MATH
MEAN
WRITING
MEAN
NUMBER OF ERRORS
TRIALS

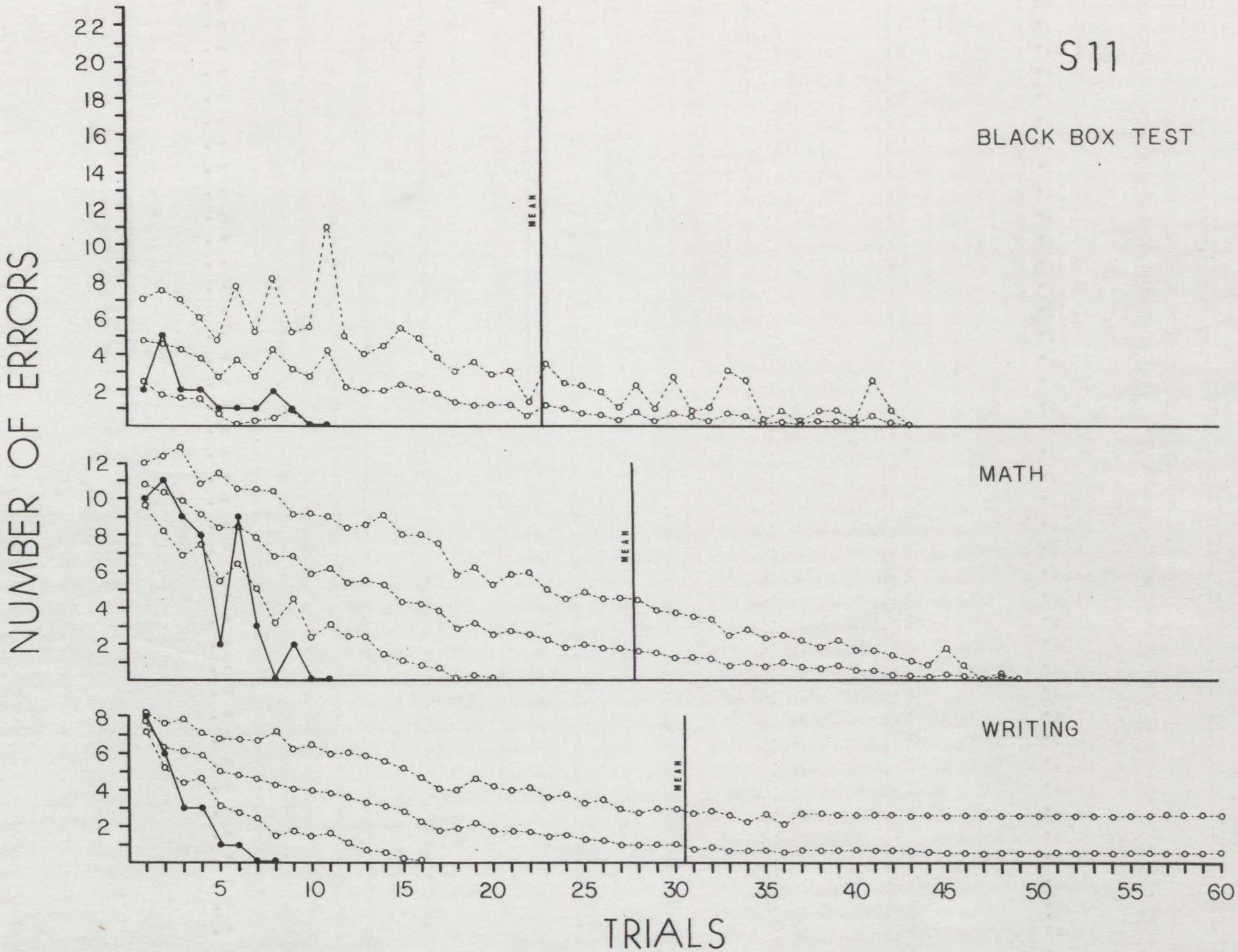

S6
BLACK BOX TEST
MATH
WRITING
NUMBER OF ERRORS
TRIALS
MEAN
22
20
18
16
14
12
10
8
6
4
2
5
10
15
20
25
30
35
40
45
50
55
60

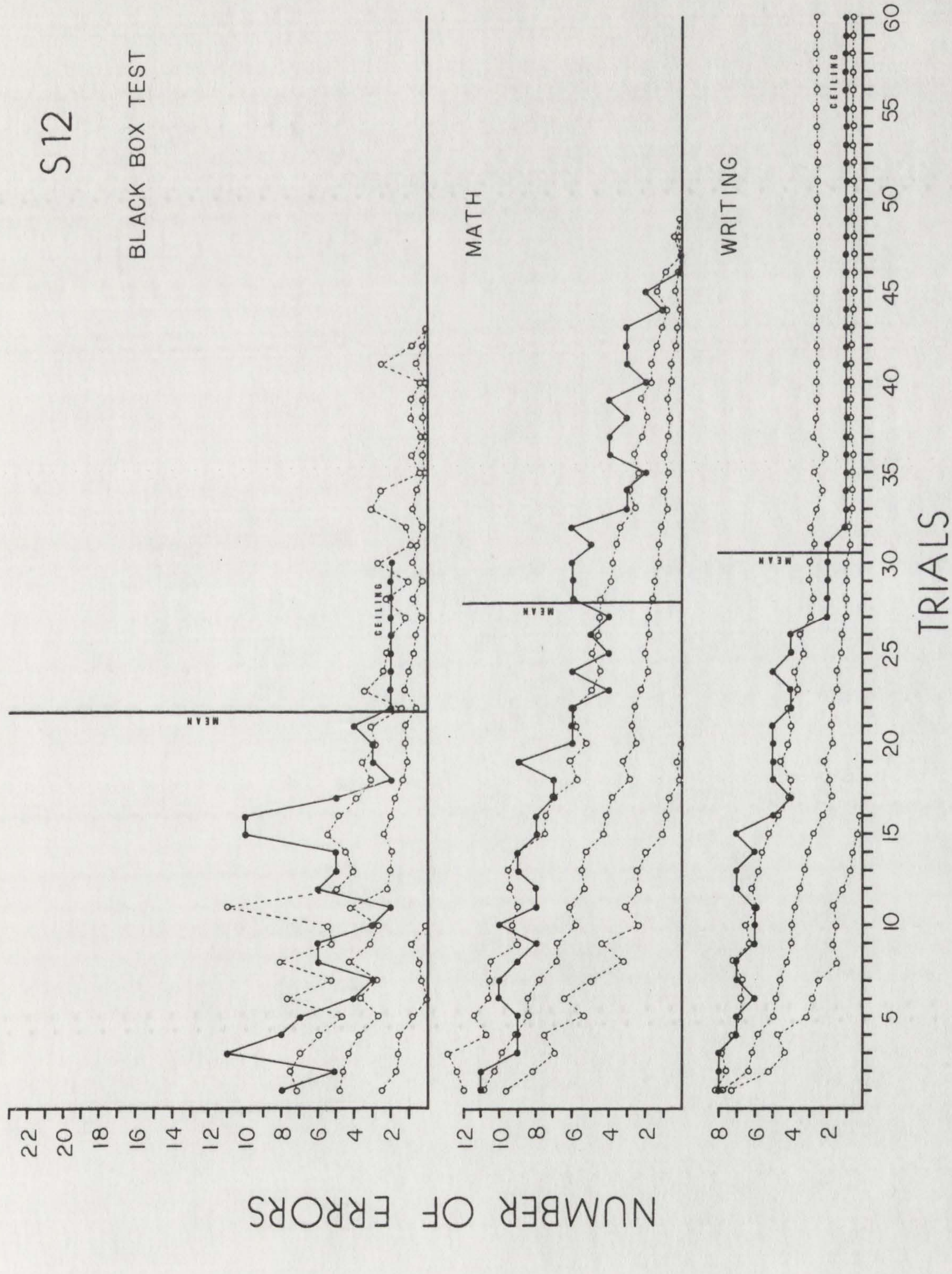

S7
BLACK BOX TEST
MATH
WRITING
CEILING
MEAN
NUMBER OF ERRORS
TRIALS


S8
BLACK BOX TEST
MATH
WRITING
MEAN
MEAN
MEAN
NUMBER OF ERRORS
TRIALS
22
20
18
16
14
12
10
8
6
4
2
12
10
8
6
4
2
8
6
4
2
5
10
15
20
25
30
35
40
45
50
55
60


S9
BLACK BOX TEST
MATH
WRITING
MEAN
MEAN
MEAN
NUMBER OF ERRORS
TRIALS


S10
BLACK BOX TEST
MATH
WRITING
NUMBER OF ERRORS
TRIALS
MEAN


S11
BLACK BOX TEST
MATH
WRITING
MEAN
NUMBER OF ERRORS
TRIALS
22
20
18
16
14
12
10
8
6
4
2
12
10
8
6
4
2
8
6
4
2
5
10
15
20
25
30
35
40
45
50
55
60


S 12
BLACK BOX TEST
MATH
WRITING
MEAN
CEILING
NUMBER OF ERRORS
TRIALS
22 20 18 16 14 12 10 8 6 4 2
12 10 8 6 4 2
8 6 4 2
5 10 15 20 25 30 35 40 45 50 55 60

VITA

ROBERT STEPHEN KNOX

Candidate for the Degree of

Doctor of Philosophy

Dissertation: A Comparison of Second Grade Children's Learning Curves on School Tasks with Their Respective Performances on the "Black Box Test of Learning Ability"

Major Field: Psychology

Biographical Information:

Personal Data: Born at Glendale, California, September 30, 1948, son of John F. and Jeannette E. Knox; married Sheila Ann Stephens, August 15, 1970; one child, Kelly.

Education: Attended elementary school in LaCanada, California; graduated from St. Francis High School in 1966; received the Bachelor of Arts degree from Gonzaga University, Spokane, Washington, with a major in psychology in 1970; completed requirements for Masters of Science in Psychology at Utah State University, Logan, in 1973; completed requirements for Doctor of Philosophy Degree in Counseling Psychology with a minor in Special Education in 1974.

Professional Experience: From September 1969 to June 1970, teacher assistant in psychology, Gonzaga University; March 1971 to June 1971, behavior modifier, Utah School for the Deaf; September 1971 to June 1972, graduate research assistant in psychology department at Utah State University; Psychometrist at UAF, Utah State University, August 1972 to June 1973; Supervisor of Clinical Services, UAF, Exceptional Child Center, USU, January 1973 to June 1973; Co-ordinator of vocational student remediation, UAF, Utah State University, January 1973 to June 1973; Vocational Counselor, USU, March 1973 to June 1973; Co-ordinator of evaluation for Title I projects (Idaho) of the Developmental Disabilities

Council, March 1973 to January 1974; Consultant to the Human Development Center, Pocatello, Idaho, June 1973 to August 1973; Supervisor of school psychology internships and practicums for Idaho State University, September 1973 to present; Psychologist for "Laboratory for the Diagnosis and Treatment of Learning Difficulties" September 1973 to June 1974; Psychologist for Pocatello School District, August 1974 to present.

Awards: Alpha Sigma Nu (National Jesuit Honors Society), April 1970; Who's Who Among Students in American Universities and Colleges, 1969 to 1970; National Student Register, 1969-1970.