DOCUMENT RESUME

ED 197 509 EC 131 711

AUTHOR Deno, Stanley L.: And Others
TITLE Relationships Among Simple Measures of Written Expression and Performance on Standardized Achievement Tests.
INSTITUTION Minnesota Univ., Minneapolis. Inst. for Research on Learning Disabilities.
SPONS AGENCY Bureau of Education for the Handicapped (DHEW/OE), Washington, D.C.
REPORT NO IRLD-RR-22
PUB DATE Jan 80
CONTRACT 300-77-0491
NOTE 109p.: For related documents, see EC 131 709-720.

EDRS PRICE MF01/PC05 Plus Postage.
DESCRIPTORS Academic Achievement: *Evaluation Methods: Exceptional Child Research: *Formative Evaluation: *Learning Disabilities: *Writing (Composition)

ABSTRACT

As a part of a project to develop formative evaluation systems for teachers to improve learning disabilities (LD) service programs in written expression, three concurrent validity studies were conducted on the relationship between performances on behavioral measures of written expression and performances on standardized achievement measures of written expression. Results of correlational analyses are reported for five behavioral measures, two standardized measures, and the Developmental Sentence Scoring System. Considerable evidence was obtained for the validity of all behavioral measures except T-units as indices of achievement in written expression. Additional comparisons indicated that both educational status (LD resource program vs regular program) and grade level were related to performance on the four validated behavioral measures. Two to 3 minute writing samples were of sufficient length to produce valid indices of written expression using the behavioral measures. Potential usefulness in monitoring student growth and evaluating the effectiveness of interventions is discussed. (Author/CL)

University of Minnesota

Research Report No. 22

RELATIONSHIPS AMONG SIMPLE MEASURES OF WRITTEN EXPRESSION AND PERFORMANCE ON STANDARDIZED ACHIEVEMENT TESTS

Stanley L. Deno, Phyllis K. Mirkin, and Douglas Marston

Institute for Research on Learning Disabilities

Director: James E. Ysseldyke

Associate Director: Phyllis K. Mirkin

The Institute for Research on Learning Disabilities is supported by a contract (300-77-0491) with the Bureau of Education for the Handicapped, Department of Health, Education, and Welfare, U.S. Office of Education, through Title VI-G of Public Law 91-230. Institute investigators are conducting research on the assessment/decision-making/intervention process as it relates to learning disabled children. Research activities are organized into eight major areas:

I. Adequacy of Norm-Referenced Data for Prediction of Success

II. Computer Simulation Research on the Assessment/Decision-making/Intervention Process

III. Comparative Research on Children Labeled LD and Children Failing Academically but not Labeled LD

IV. Surveys on In-the-Field Assessment, Decision Making, and Intervention

V. Ethological Research on Placement Team Decision Making

VI. Bias Following Assessment

VII. Reliability and Validity of Formative Evaluation Procedures

VIII. Data-Utilization Systems in Instructional Programming

Additional information on these research areas may be obtained by writing to the Editor at the Institute.

Research Report No. 22

RELATIONSHIPS AMONG SIMPLE MEASURES OF WRITTEN EXPRESSION AND PERFORMANCE ON STANDARDIZED ACHIEVEMENT TESTS

Stanley L. Deno, Phyllis K. Mirkin, and Douglas Marston

Institute for Research on Learning Disabilities

University of Minnesota

January, 1980

Table of Contents

Abstract

Three concurrent validity studies were conducted on the relationship between performances on behavioral measures of written expression and performances on standardized achievement measures of written expression. Results of correlational analyses are reported for five behavioral measures, two standardized measures, and the Developmental Sentence Scoring System. Considerable evidence was obtained for the validity of all behavioral measures except T-units as indices of achievement in written expression. Additional comparisons indicated that both educational status (LD resource program vs. regular program) and grade level were related to performance on the four validated behavioral measures. Two to three minute writing samples were found to be of sufficient length to produce valid indices of written expression using the behavioral measures. The potential usefulness of these measures to monitor student growth and evaluate the effectiveness of interventions is discussed.

Introduction

The research reported here was conducted as part of a project that has as its purpose developing formative evaluation systems for teachers to use in improving learning disabilities service programs in written expression. The primary assumptions upon which that research project is based are:

1) that the success of learning disabilities services is defined primarily by the extent to which those services improve the academic and social behavior goals of the individual students served,
2) that teachers can increase the success of learning disabilities services by systematically measuring student progress toward achievement of program goals and then adjusting student programs to improve that progress, and
3) that the technology presently available for teachers to use in measuring student progress and adjusting programs based on measured progress is either not sufficient or has not been sufficiently tested.

The particular part of the research project described here was conducted to answer a first and critical question that is raised when developing a formative evaluation system in written expression: what student performance data can be routinely and easily obtained that validly index achievement? The question arises because, for several good reasons, commercially prepared standardized tests of written

expression ordinarily used to assess achievement cannot be used routinely in a formative evaluation system to monitor performance. First, commercially produced standardized tests take too much time to administer. Second, an insufficient number of equivalent forms is available for any test to be used in the repeated measurement of performance required for formative evaluation. Third, the cost of using achievement tests repeatedly is prohibitive.

The development of measurement procedures that can be incorporated relatively easily into the daily routine of most teachers working in learning disabilities programs is deemed desirable if intensive monitoring of program effects on student performance is to occur. The importance of intensively monitoring program effects is that such monitoring enables us to more precisely determine the appropriateness of services provided to individual students. Given the requirement in P.L. 94-142 (Federal Register, 1977) that each handicapped student be provided an "appropriate educational program" and our current inability to diagnose and prescribe effective programs (Arter & Jenkins, 1978), continuous evaluation of a student's program is the only way to achieve substantive compliance with the law (Deno & Mirkin, 1980).

Beyond compliance with the law, research on the use of intensive repeated measurement in formative evaluation of instruction has already yielded evidence bearing on its potential benefits (Bohannon, 1975; Crutcher & Hofmeister, 1975; Frumess, 1973; Lovitt, Schaff, & Sayre, 1970; Mirkin & Deno, 1979). The research findings are isolated, however. The research and development program of which the present studies are a part was designed to systematically construct formative evaluation

procedures for learning disabilities programs that specify:

(1) What behaviors to measure when improved writing proficiency is an IEP goal.

(2) How to repeatedly measure those behaviors reliably.

(3) Who should administer the measurement procedures.

(4) How often measurement should occur.

(5) How to obtain data most efficiently.

(6) How to use repeated measurements of student performance to increase intervention effectiveness.

The strategy employed in the present research was first to review available literature on written expression, to identify behaviors that are commonly used to assess achievement in this academic domain; second, to develop measurement procedures for taking data on those behaviors; and third, to determine the reliability and validity of the measures by correlating the scores obtained on them with scores from standardized measures of written expression that are highly respected, and technically adequate with respect to their psychometric properties.

To be considered for inclusion in a formative evaluation system the developed measures had to fulfill the following criteria:

(1) They must be valid with respect to widely used measures of achievement in written expression.

(2) They must be immediately sensitive to the effects of relatively small adjustments made in (a) instructional methods and materials, (b) motivational techniques, and (c) administrative arrangements (e.g., adjustments in grouping, setting for instruction, teacher/tutor,

time of instruction, etc.).

(3) They must be <u>easy to administer</u> by teachers, parents, and students.

(4) They must include many parallel forms that are <u>frequently administrable</u> (daily, if necessary) to the same student.

(5) They must be <u>time efficient</u>.

(6) They must be <u>inexpensive</u> to produce.

(7) They must be <u>unobtrusive</u> with respect to routine instruction.

(8) They must be <u>simple to teach</u> to teachers, parents, and children.

Our hope is that regardless of personal philosophical, theoretical, historical, and current situational constraints, those responsible for ensuring the quality of learning disabilities services will continuously evaluate the impact of those services on the academic and social behaviors of their individual students. The measurement procedures that are described here are an important first step in the development of such an evaluation system.

Review of the Literature

A wide array of approaches to measuring a child's performance in written discourse is currently available. The methods range from formal achievement testing, to analysis of grammatical and syntactic structures, to counting such basic written units as words. In this review of the literature on the measurement of written expression, these approaches will be described briefly and appraised for their appropriateness in formative evaluation procedures.

Aside from the traditional assignment of letter grades to written essays, achievement testing is probably the most popular approach to assessment of writing performance. Tests such as the Stanford Achievement Test (Madden, Gardner, Rudman, Karlsen, & Merwin, 1978), SRA Achievement Series: Language Arts (Thorpe, Lefever, & Naslund, 1968), and the Metropolitan Achievement Tests (Durost, Bixler, Wrightstone, Prescott, & Balow, 1971) employ varying subtests to evaluate a student's "language ability," "punctuation," "spelling," "grammatical structure," "sentence sense," "word usage," and "capitalization." The format of achievement test items is usually forced-choice, often requiring multiple-choice responses, and at times, completion or underlining. Although the achievement tests are often well standardized, with reported validity coefficients and sufficient reliability, they do not meet many of the requirements of a formative evaluation system which entails the frequent monitoring of a student's progress.

Several limiting factors make standardized achievement tests inappropriate for formative evaluation of written expression. First, the behavior sample in these achievement tests is limited and is dependent on reading

skills. In fact, it could be argued that many of the achievement tests that purport to measure "written expression" primarily measure reading skills rather than language development or writing skills. A fairly new achievement test that does employ a written production task is Hammill and Larsen's (1978) Test of Written Language (TOWL). Their incorporation of the student's written sample in the assessment of writing skills appears to be a promising alternative to forced-choice items.

A second limiting factor in using achievement tests for formative evaluation is the duration of the tests. Achievement tests clearly are designed for infrequent administration, with many taking from 30 to 60 minutes for completion.

Third, item samples on the various subtests are often small, reducing the likelihood that the subtests will be sensitive to changes in the individual's performance. As stated before, formative evaluation systems must be sensitive to change in performance.

There have been numerous unique attempts to measure written expression that do not fall into the achievement test domain. Hunt's (1965, 1966) Minimal Terminable Unit or T-unit length, for example, is a measure of syntactic complexity. A growing body of research demonstrates that the T-unit is a valid measure of written expression (Perron, 1977), is reliable (Dixon, 1972), and increases with age (O'Donnell, Griffin, & Norris, 1967; Veal, 1974). Initially, the T-unit length appears to be a promising candidate for use in a formative evaluation system. A major problem with T-unit length, however, is its probable insensitivity to change. Since the average T-unit length ranges from only one word in Grade 1 to 15 words in Grade 12, weekly or monthly

changes in performance would be difficult to discern.

Related to Hunt's T-units are Dixon's (1972) free modifiers and Golub and Kidder's (1974) Syntactic Density Score. The latter measure incorporates words per T-unit, subordinate clauses, words per main clause, number of modals, number of prepositional phrases, and number of possessive nouns and pronouns into an equation that yields a performance score. Dixon's free modifiers and the Syntactic Density Score both correlate highly with criterion measures of written expression. However, both are difficult to score.

The Developmental Sentence Scoring System (Lee & Canter, 1971), originally created to measure syntactic acquisition and the child's knowledge of grammatical rules in oral language, has been adapted for indexing written discourse. The scoring procedures consist of eight scales that measure the student's use of noun modifiers, personal pronouns, main verbs, secondary verbs, negatives, conjunctions, interrogative reversals, and WH-questions. Rubin, Buium, and Balow (1975) analyzed the grammatical sophistication of written compositions of nine-year-old children with this measure and found that the index generally followed a developmental continuum. Despite the measure's appeal, however, the complex scoring system makes it much too laborious for use in a formative evaluation system.

A wide variety of subjective measuring scales also have been developed to measure performance in writing (Diederich, 1964; Mullis, 1976; Stahl, 1974). The scales are subjective in that, without detailing specific criteria, the judge is asked to rate students' written compositions. In the Primary Trait Scoring System (Mullis, 1976), for example, the

judge must rate the student's composition on a scale from 1 to 4 on the categories of "fantasy," "insights," "entry world of picture," "dialogue," "point of view," and "tense." The Diederich Scale (1964) is similar in that the judge rates the writer as low, middle, or high on "quality and development of ideas," "organization," "style," "wording," "grammar," "punctuation," "spelling, and "handwriting." Likewise, Stahl's (1974) Feature Analysis rates the writer on "order," "principle of selection," "syntax," "balance," "organization," "connectives," and "openings." In all, the subjective scales may be praised for their attempt to measure those dimensions of written discourse that are difficult to quantify precisely. However, the scales may not serve well to continously monitor student performance, since validity and reliability data are difficult to obtain for ratings that are so ambiguously defined.

Carroll, Davies, and Richman (1971) have provided a unique approach to evaluating writing skills that involves determining the maturity of the student's word choices. These authors developed the Standard Frequency Index as a measure of mature word usage. For Carroll et al., the frequency with which a word appears in the English language is inversely related to its maturity. Thus, the use of a word that is not commonly used connotes maturity. Using the Standard Frequency Index, Finn (1977) developed a list of 220 words that do not discriminate fourth grade from eleventh grade students on mature word usage. Finn refers to this group of words as the Undistinguished Word Choice List. While scoring compositions in terms of word maturity looks promising, it remains to be seen whether such a measure can be reliably and quickly scored, is sensitive

to growth, and relates to other measures of written expression. Page (1968) does provide some evidence for this approach. He found that words from high school compositions that were also found on the Dale word list (Dale & Chall, 1948) were inversely correlated with performance.

Another unique way of examining written performance is to document the number of large words included in the composition. Page (1968) reported a moderate correlation between the average word length measured by letters and criterion measurement of written performance. Kuehnle, Tindal, and Marston (1979) examined the compositions of 20 third-grade students and found that large words (seven or more letters) correlated .47 with creativity.

Cloze procedures also have been used to measure written discourse (Nystrand, 1977). Nystrand argues that this procedure better assesses individual strengths and weaknesses than T-units and allows one to judge the syntactic, semantic, and graphic dimensions of writing. Extensive empirical support for this argument is absent, however.

Research also indicates that spelling is a significant component of written expression (Page, 1968; Slotnick, 1972). Page (1968) found that the number of spelling errors was inversely correlated to performance. Slotnick's (1972) factor anaysis of compositions written by high school students determined that the second most important factor was spelling.

Myklebust (1965) developed a test of written language that employs a variety of measures of writing. His Picture Story Language Test evaluates the writer on three dimensions: productivity, syntax, and quality

of content. Although the specific measures on all three dimensions appear to index development of writing skills between the ages of nine and 15, it is total words, a productivity measure, that may offer most to a formative evaluation model. This measure is obtained easily and can range widely across ages. Total words written, in addition, appears to be the most powerful discriminator between children with learning disabilities and those who develop normally (Myklebust, 1965). Hillerich (1971) agrees that total words is a potent measure and states that fluency is an excellent indicator of development in written language. The research of Brigham, Graubard, and Stans (1972) also substantiates this claim. These authors note that number of words written, number of different words written, and number of new words written, all index growth in written discourse. Rubin and Buium (1974) found that total words written correlated highly with abstractiveness or creativity and vocabulary richness. Slotnick (1972) factor analyzed compositions written by high school students and identified fluency or total words written as the most significant factor contributing to written performance, accounting for 24 percent of the variance. Page (1968) also reported a moderate correlation between total words written and written performance.

As evidenced here, there are many available approaches for measuring written expression. For any of these approaches to be employed in a formative evaluation system, they must first satisfy the suggested criteria. The literature reviewed here suggests five measures of written expression that might be used: Hunt's mean T-unit length, Myklebust's total words written in a composition, number of Large Words used, number of words spelled correctly, and Finn's Undistinguished word list.

Three Studies of Written Expression

Three studies were conducted to identify behavioral units in written expression that could be measured in a formative evaluation system and to determine the concurrent validity of these measures. General information on the three studies is presented first, followed by the specific methods and results of each study.

General Procedures

When the criteria established for inclusion of a measure in a formative evaluation system were initially applied to those behaviors included in the literature on written discourse, five behavioral units were identified as potentially useful for a formative system. The measures were: T-units, words written, mature words written, large words written, and words spelled correctly in a written sample.

Subjects in the three studies were asked to write compositions in response to a picture stimulus, story starter, or topic sentence. The subjects' written samples were then scored to obtain measures of the five behaviors. To establish concurrent validity, standardized achievement tests of written expression were administered also, and performance of students in learning disability resource programs was contrasted to that of regular class students. Scores for the five behavioral units were then correlated with performance on the criterion achievement measures.

Stimulus Materials

Three procedures were employed to elicit written samples from the subjects: a picture stimulus, story starters, and topic sentences. The Picture Stimulus procedure closely followed Myklebust's (1965) Picture

Story Language Test. Every effort was made to adopt his criteria. In the course of the study two pictures were used, picture X-3 from the Peabody Language Development Kit, Level 3 (Dunn & Smith, 1967), and the picture stimulus from Hammill and Larsen's (1978) Test of Written Language (TOWL). The stimulus picture from the Peabody Kit is a 21.5" by 18.5" colored picture depicting a young boy riding an elephant in the jungle accompanied by many different types of animals. In the background a fire rages as a helicopter hovers over it. The picture from the TOWL is a 3-picture sequence showing people on earth being visited by flying saucers, flying through space, then landing on a different planet.

Other methods for eliciting children's writing responses included Story Starters and Topic Sentences. In the Story Starter condition, a written stimulus was shown and read aloud to the child. The stimulus sentence created a situation in which the child was asked to imagine himself or herself. The child was then asked to write a story about what happened. Topic sentences generally were shorter written stimuli for which the child was asked only to "write something about this." For all three stimuli the child was asked to write for five minutes. Examples of story starters and topic sentences can be found in Appendix A.

Scoring Procedures

Each student's written sample was scored to obtain measures of the five behavioral units.

T-unit length. Scoring was accomplished according to Hunt's rules (1965, 1966). An example of how T-unit length was computed for written compositions is provided in Appendix B.

Mature Words. Scoring procedures were based on Finn's (1977) Undistinguished Word Choice List. Words written that did not appear on that list were circled and totaled. Appendix C shows how compositions were scored for mature words.

Total Words Written. Scoring involved simply counting the number of words written in the sample. Words did not have to be spelled correctly. The minimum criterion for inclusion of a word was that at least two letters be written in sequence, except for single letter words (e.g., "I" and "a"). Appendix D provides an example of scoring on this measure.

Large Words. Scoring involved counting the number of words with seven or more letters. In later phases of research, large words ending in "ed" or "ing" were counted only if the root word had seven or more letters. An illustration of such scoring may be found in Appendix E.

Words Spelled Correctly. Scoring written samples for spelling was done by checking those words spelled incorrectly and tallying the remaining words. An example of this measure is provided in Appendix F.

Criterion Measures

To establish the concurrent validity of the measures of written expression, nine different criterion variables were used individually and in combination.

Test of Written Language (TOWL). Developed by Hammill and Larsen (1978), the TOWL was used in all three phases of investigation. The TOWL consists of five subtests: Vocabulary, Thematic Maturity, Spelling, Word Usage, and Style. The criterion measures obtained from the TOWL were the raw scores from the five subtests and the raw total of the

five subtests. It should be noted that four of the five subtests are 25-item scales, and therefore were equally weighted in the raw total. However, the Vocabulary subtest does not have a limited number of items and as a result students may obtain a score higher than 25 points on this subtest. We found the range of Vocabulary scores in our sample of 135 students to be from zero to 70 points. To ensure that the Vocabulary subtest was equally weighted in the raw score total, a student's score on this measure was multiplied by a correction factor of .357 (or 25 ÷ 70). A Written Language Quotient (WLQ) with an age-scaled total score mean of 100 can be derived from the subtests and was also used as a criterion measure.

Stanford Achievement Test, Intermediate I, Word Usage Subtest. This subtest, constructed by Madden et al. (1978) was also administered as a criterion measure to the subjects. The Word Usage subtest includes 42 multiple-choice items designed to measure appropriate choice of words within written context.

Developmental Sentence Scoring System. The Developmental Sentence Scoring System (Lee & Canter, 1971) was developed originally to describe syntax in oral language. In the present study it was used as a criterion measure of the syntactic maturity of the written samples. Appendix G outlines the scoring categories adopted from this approach.

Program Placement. Since a primary goal of the research was to develop measures useful for assessing students receiving learning disabilities services, one validity criterion is the degree to which the measures discriminate between students receiving LD services and those

not receiving such services. To this end, the performance of random samples of students from regular classrooms not receiving special education services was compared to the performance of students in LD resource programs on all measures.

Study I

Subjects

Participation was solicited from two public elementary schools in Minneapolis. Twenty-eight children were randomly selected from those schools. The children were enrolled in grades three through six and ranged in age from seven to 11 years. Half were boys and half were girls. Twelve of the 28 children were receiving learning disabilities services in resource programs for approximately one hour per day. Sixteen children were in regular classes all day.

Procedure

Testing was conducted by three graduate research assistants, trained at the Institute for Research on Learning Disabilities at the University of Minnesota. Each child was tested individually and data collection required approximately 55 minutes per child.

Each child was presented a Story Starter developed by Institute staff and the picture stimulus from the Peabody Language Development Kit. A copy of the Story Starter may be found in Appendix H.

For the Picture Stimulus condition the examiner presented the student with the response form and two pencils. He or she then read the following instructions:

> Today I want you to write a story. First, I want you to look at this picture and think about what is happening in

> it. Then I want you to write a story about what is happening in this picture. You will have a minute to think about the story you will write and then have five minutes to write it. When I say, "please start writing," you may begin.

For the Story Starter condition the instructions were similar. The examiner recited the following instructions:

> I want you to write another story. I am going to read a sentence to you first, and then I want you to write a short story about what happens. You will have a minute to think about the story you will write and then have five minutes to write it. When I say "please start writing," you may begin.

In both writing situations the student was given five minutes to finish his or her composition. In addition, the examiner recorded the amount of time the student actually spent writing. This procedure allowed examination of the measures as rate variables.

All students also were administered the TOWL. Administration procedures outlined in the TOWL's instructional manual were followed closely. The duration of the test was approximately 40 minutes.

Results

All information collected was analyzed by computer at the University of Minnesota Computer Center. Statistical Package for Social Sciences (SPSS) was used exclusively for all statistical analysis.

Analyses of the Study I results focused on the correlations between scores for the five behavioral units in written expression and scores derived from the TOWL subtests. A single score was obtained for each of four of the behavioral units and two scores were obtained for Words Written - total words written, and rate. This latter score was created by dividing the total number of words written by the number of minutes each student wrote during the given five minute interval.

Picture Stimulus. The correlational results for the measures obtained from the picture stimulus are presented in Table 1. As inspection of Table 1 reveals, the highest correlations between scores for the behavioral units and scores on the TOWL were found for Words Spelled Correctly (.38 to .88), Total Words Written (.65 to .82) and Mature Words (.56 to .79). Mean T-unit length generally was not associated with performance on the TOWL, while correlations between rate of words written and the TOWL were relatively low.

- - - - - - - - - - - - - - - - - -

Insert Table 1 about here

- - - - - - - - - - - - - - - - - -

The discrepancy obtained in correlations of Total Words Written and Rate of Words Written with the TOWL is of interest. Since the content of the two measures is highly similar, one might expect approximately the same correlations with the criterion achievement measure. Further analysis, however, revealed that lower performing students would often write for only a few seconds and then quit. The result was that their rate of performance was similar to students who had scored high on the achievement measures. It appears that, in this situation, the rate measure masked the differences between low and high functioning children. Total Words Written seems to provide a finer discrimination of writing skills.

Story Starters. An analysis of the relationships between scores on written compositions cued by Story Starters and the criterion measures was also conducted. These results are presented in Table 2.

- - - - - - - - - - - - - - - - - -

Insert Table 2 about here

- - - - - - - - - - - - - - - - - -

The pattern of correlations obtained with the story starter was quite similar to that obtained with the picture stimulus. The obtained coefficients for Mean T-unit length and Large Words appear to have increased slightly; however, the coefficients for Mean T-unit length (.11 to .26) are still low. The obtained correlation coefficients for the other measures are smaller than in the Picture Stimulus condition, with correlations for Words Spelled Correctly ranging from .45 to .73, Total Words Written ranging from .41 to .70, Mature Words ranging from .41 to .71, and Large Words ranging from .45 to .72.

In all, the results of Study I offered preliminary indications of the validity of some of the formative measures and provided impetus for the further investigation of these measures in Study II.

Study II

In addition to replicating Study I, Study II was designed to investigate three questions important to the development of the formative measures. First, to examine the possibility that the results of Study I were specific to the Story Starter used, student performances on three independent story starters and topic sentences were analyzed to determine whether student performance on these samples were significantly different. Second, to further explore the types of writing stimuli that might be used as writing cues, the Topic Sentence was included as a test stimulus. Third, to ascertain the writing time necessary to obtain a technically adequate measure, the performance of subjects at varying time intervals was analyzed.

Subjects

Twenty-eight children were randomly selected from an elementary school in the Twin Cities metropolitan area. The sample was composed of 16 males and 12 females, ranging in grade placement from grade three to grade six, and in age from eight to 12 years. Four of the 28 children were receiving learning disabilities services in a resource program for approximately one hour per day; the remaining 24 children were in regular classes all day.

Procedure

Two graduate research assistants and two undergraduate students from the College of Education at the University were trained to administer the tests. Each child was tested in a small group of three or four students. Data collection lasted approximately 55 minutes per group.

Each child was given a Story Starter that had been randomly selected from a group of three Story Starters. The picture stimulus from the Peabody Language Development Kit was also used with all students. In addition, the students were randomly assigned to write about one of three Topic Sentences. Children had five minutes to write compositions in each of the three stimulus conditions. On all three compositions, the examiner determined how far the student had written at the end of one minute, two minutes, three minutes, and four minutes. Copies of the Story Starter and Topic Sentences employed in this study may be found in Appendix I.

The TOWL was used as one of the validation criteria. Subtest 8 of the language section of the Stanford Achievement Test, Intermediate I (Madden et al., 1978) also was administered to all subjects. Except

for the additional procedure of marking one minute intervals on the written compositions, instructions to the students were identical to Study I specifications. Order of presentation of stimulus materials is outlined in Appendix J.

Scores for the written samples were obtained using the same scoring procedures as in Study I. An additional modified Large Word score was obtained by omitting words with "ed" and "ing" endings in the scoring of large words. Root words had to contain seven or more letters to be counted as large words.

Results

Story Starters. To answer the question as to whether students perform inconsistently on different Story Starters, a one-way analysis of variance was conducted. Results indicated that there were no significant differences among the three story starters for Total Words Written (F = 1.1, p = .348), Mature Words Written (F = .229, p = .797), and Words Spelled Correctly (F = .536, p = .592).

Correlation coefficients obtained when using a Story Starter stimulus are presented in Table 3. The correlations obtained across cumulative minutes of writing are included in the table. As can be seen, the obtained coefficients between scores from the writing sample and Total Raw Score on the TOWL parallel Study I correlations, except that the coefficient for Large Words is somewhat higher. Of related interest is the fact that little difference was obtained between the coefficients for Large Words and Large Words Modified. The coefficients for Mean T-unit Length are higher, and in contrast to Study I, they are significant. Still, the highest coefficients obtained are for Mature Words

(.82), Total Words (.79), and Words Spelled Correctly (.78).

- - - - - - - - - - - - - - - - - -

Insert Table 3 about here

- - - - - - - - - - - - - - - - - -

The Stanford Word Usage Subtest correlations provide a picture simular to those obtained from the TOWL. The relative size of the coefficients for the different scoring procedures remains essentially the same. Coefficients with the Stanford Word Usage Subtest are lower than with the TOWL in every case except for Large Words Modified, where they are somewhat higher.

Inspection of Table 3 with respect to the question of sample duration reveals mean correlations that range from .60 for a one minute writing sample to .67 for a five minute sample. Sample duration appears to influence coefficient size, but only slightly. The increase in coefficients from .60 at one minute to .65 at three minutes may be practically important, but little difference exists among coefficients for three, four, and five minute samples.

Topic Sentence. To determine whether significant differences exist among various Topic Sentences, a one way analysis of variance was conducted. Again, the results indicated no significant differences among Topic Sentences 1, 2, and 3 for the various measures: Total Words Written ($F = .26$, $p = .775$), Mature Words Written ($F = .45$, $p = .644$), and Words Spelled Correctly ($F = .32$, $p = .728$). As may be seen in Table 4, the pattern of correlations for Topic Sentences is similar to the pattern of coefficients presented for Story Starters. Again, Total Words Written, Mature Words, and Words Spelled Correctly offer

the highest coefficients. For the duration data, the coefficients are virtually identical to the Story Starter correlations. Again, there is essentially no difference between three, four, and five minute samples.

- - - - - - - - - - - - - - - - - -

Insert Table 4 about here

- - - - - - - - - - - - - - - - - -

Picture Stimulus. Correlation coefficients for the Picture Stimulus condition (see Table 5) were slightly higher when compared to Story Starters and Topic Sentence data, for all formative measures of written expression. Nonetheless, the pattern of correlations is similar. Mature Words, Words Spelled Correctly, Total Words Written, as well as Large Words, showed high correlations (above .70) with both the TOWL and the Stanford. Again, there appeared to be little difference between written samples of three, four, and five minute durations.

- - - - - - - - - - - - - - - - - -

Insert Table 5 about here

- - - - - - - - - - - - - - - - - -

Relationships among the various writing formats. To compare performance among the different stimuli (i.e., Story Starter, Topic Sentence, and Picture Stimulus), an intercorrelation matrix was generated. The results, presented in Table 6, suggested a high degree of correspondence between Story Starters and Topic Sentences, Story Starters and Picture Stimulus, and Topic Sentences and Picture Stimulus, on all formative measures except Large Words written.

- - - - - - - - - - - - - - - - - -

Insert Table 6 about here

- - - - - - - - - - - - - - - - - -

However, high correlations cannot be interpreted to mean that the performance of students across the three measures is equivalent. This question may be resolved by examining the mean performance of the 28 elementary students on the three formative measures. These scores, presented in Table 7, reveal little difference in the performance of students when the type of written stimulus is considered. To substantiate this claim, a repeated measures analysis of variance was computed for each of the three formative measures in Table 7. The results indicated no significant differences in performance on Total Words Written (F = .98, p = .381), Mature Words (F = .78, p = .465), and Words Spelled Correctly (F = .80, p = .456) across the three stimulus formats.

- - - - - - - - - - - - - - - - - -

Insert Table 7 about here

- - - - - - - - - - - - - - - - - -

Study III

Subjects

Eighty-two children were randomly selected from five elementary schools in the Twin Cities area. The 42 males and 40 females were in grades three through six and ranged from seven to 11 years old. Thirty-one of these children were in learning disability resource programs.

Procedure

Study III was a replication of Study II with two modifications. First, the Developmental Sentence Scoring System (DSS) was employed, in addition to the TOWL, to provide supplementary evidence for validation.

The DSS can be used to measure syntactic maturity and, while too complex for routine use in formative evaluation, was useful as a criterion for establishing concurrent validity. Sample size for Study III was increased substantially as well, to ensure reliability of the findings.

Procedures were again similar to those in Studies I and II. Subjects were presented a Story Starter, Topic Sentence, and a Picture Stimulus and given five minutes in which to write a composition. Copies of the Story Starter and Topic Sentence may be found in Appendix K. Examiners also noted the subject's progress at the end of each minute.

Results

Scoring procedures for the written samples were the same as for Study II. In addition, samples were scored for number of correct Letter Sequences. While several of the other scoring procedures apparently yield valid data, they may not be sensitive to growth in low functioning students in learning disabilities program who may only be writing letters. An illustration of how this measure was used to score a written composition is given in Appendix K.

Since Study II results revealed no clear differences in the pattern of correlation coefficients among the three types of stimuli, only the results for writing samples obtained using Story Starters are presented here. Table 8 provides the validity data with respect to the DSS. Analyses that include the TOWL scores are reported in the section where data from all three studies are combined. The correlation coefficients in Table 8 reveal that the production measures (i.e., Total Words Written, Words Spelled Correctly, and Letters in Correct Sequence) correlated highest with the criterion of syntactic maturity as measured

by the DSS. The coefficients for types of words (i.e., Large Words, Large Words Modified, and Mature Words) were lower than in previous studies particularly for Large Words and Large Words Modified. Mean T-unit length also was quite low when compared to the other formative measures.

- - - - - - - - - - - - - - - - -

Insert Table 8 about here

- - - - - - - - - - - - - - - - -

Analysis of the data for different sample durations again revealed little difference among three, four, and five minute writing samples. The coefficients for one minute samples were substantially lower.

Combined Sample Results

To gain power in the statistical analysis, data from the three studies were pooled where possible. Combining sample data was possible for three reasons. First, all subjects were randomly selected from the same grades; second, all subjects wrote a five-minute composition from a Story Starter; and third, all subjects were administered the entire Test of Written Language. A particular advantage gained from combining the sample data was the opportunity it afforded to contrast the performance of students in learning disabilities programs with regular class students. In addition, pooling the data enabled more reliable estimates of grade level performance.

Table 9 summarizes correlations among the dependent measures and the subtests of the TOWL. The relative strengths of the correlations between the scores from the writing samples remained the same for the combined sample as for the separate studies. Large Word coefficients were consistently smaller than those for the other scoring procedures. However, the magnitude of all coefficients appeared to be slightly lower, perhaps as a function of technical problems in pooling and analyzing data across the three studies. The correlational data from the combined samples substantiate the validity of Total Words Written, Mature Words, Words Spelled Correctly, and Letters in Correct Sequence as measures of written expression.

- - - - - - - - - - - - - - - - -

Insert Table 9 about here

- - - - - - - - - - - - - - - - -

To provide additional clarity concerning the relationships between scores from the writing sample and the total raw score on the TOWL, scat-

tergrams were created for Mature Words, Total Words, Words Spelled Correctly, and Correct Letter Sequences. The scattergrams are presented in Figures 1, 2, 3, and 4.

- - - - - - - - - - - - - - - - - - -

Insert Figures 1 - 4 about here

- - - - - - - - - - - - - - - - - - -

Inspection of the scattergrams reveals a functional ceiling on the TOWL at a raw score of approximately 80. No such ceiling was apparent on the scores for the writing sample. A floor does not exist for any scores except, perhaps, Mature Words. On technical grounds the ceiling on the TOWL is important because it may attentuate the correlations and explain the somewhat higher correlations between scores for written samples and the DSS obtained in Study III.

LD Program Placement as a Criterion

A particularly important criterion for establishing the concurrent validity of the different formative scoring procedures for use in the formative evaluation of learning disabilities is the extent to which the scores are consistent with program placement. Table 10 presents comparative data for both groups. Validity was examined with respect to mean differences between groups. As can be seen, mean differences existed between the groups by factors ranging from 1.5 to 2.0 times for Total Words Written, Mature Words, Words Spelled Correctly, and Letters in Correct Sequence. All differences were statistically significant, demonstrating their power to reliably discriminate between children from regular and resource classrooms. Besides supplying additional evidence of internal validity for the dependent measures, these data suggested that the formative evaluation measures might be employed successfully in the

assessment of children with learning disabilities.

- - - - - - - - - - - - - - - - -

Insert Table 10 about here

- - - - - - - - - - - - - - - - -

A question of related interest is how samples from regular and LD programs compare with respect to the pattern of correlations between the various formative measures and performance on the TOWL. Those correlations are contained in Table 11.

- - - - - - - - - - - - - - - - -

Insert Table 11 about here

- - - - - - - - - - - - - - - - -

The most obvious difference between samples was in the relative strength of the correlations. In every subtest comparison, the obtained coefficient for the LD program sample was higher than that for the regular class sample. This finding held for the correlations with Total Score on the TOWL as well. The most evident pattern in the correlations suggested that letter and word production was a better discriminator for the LD program sample than for the regular program sample.

Interpretation of the comparative correlation patterns is difficult. By extracting the LD program sample from the total, the distributions of scores are radically altered. As is illustrated in Table 12, students samples from Grades 4, 5, and 6 did not differ as much from one another as they did from students in Grade 3.

- - - - - - - - - - - - - - - - -

Insert Table 12 about here

- - - - - - - - - - - - - - - - -

When the data in Tables 13 through 16 are examined, it becomes clear that the relative difference in performance between third and

sixth grade students was much greater for the LD program sample than for the regular program sample. Further, on each of the scores from the writing sample, Grade 6 students in LD programs scored almost identically to Grade 3 regular class students.

- - - - - - - - - - - - - - - - - - - -

Insert Tables 13 - 16 about here

- - - - - - - - - - - - - - - - - - - -

One way ANOVAs were conducted to determine whether the differences among student performance for the various grades were significant. The main effects for grade level were significant for all scores from the writing samples. These results are given in Tables A to D in Appendix L.

Two way ANOVAs were conducted to examine the joint contribution of grade level and educational status to the total variance of scores. For the dependent measure Total Words Written, both main effects were highly significant (grade: $F = 10.4$, $p = .001$; educational status: $F = 23.1$, $p = .001$). The interaction was nonsignificant ($F = 1.0$, $p = .359$), suggesting that each factor's contribution to the variation in Total Words Written was unique.

The measures of Words Spelled Correctly, Mature Words, and Letters in Correct Sequence behaved similarly, with significant main effects for grade level and educational status and no interaction effect. These results are provided in Tables E to H of Appendix L.

The results of these analyses indicated that both educational status and grade level were uniquely related to performance on the behavioral measures. This suggests that the measures of written expression developed for formative evaluation accurately reflect the true level of performance in this academic area.

Discussion

The results of the research presented here provide an ample empirical base for constructing measures to be used in the formative evaluation of learning disabilities interventions designed to improve written expression. The answers to the questions raised earlier concerning what behaviors to measure and how these behaviors can be measured routinely, validly, and efficiently, seem straightforward. Among the behavioral units investigated, Total Words Written, Words Spelled Correctly, Correct Letter Sequences, and Mature Words appear to be strongly and consistently related to the criteria used - achievement tests, Developmental Sentence Scoring, and program placement.

The validities were consistent, regardless of whether pictorial or verbal stimuli were used to cue writing. The implication is that either story starters or topic sentences might well be used because of their simplicity and cost efficiency.

Time allowed for writing in response to stimulus presentation need not exceed three minutes. In fact, two minutes might be sufficient if the measures are used as often as daily.

Since counting words written is the most time efficient scoring procedure, recording that datum is recommended for students who are writing words on cue. For beginning writers, counting letters in correct sequence may be an appropriate alternative.

In sum, a daily 2-3 minute sample of the number of words written by a student in response to a story starter or topic sentence reliably discriminates between good and poor writers both in and out of learning

disabilities programs. While the measure seems overly simplistic and inelegant in contrast to the more complex measures of written expression often used, the economy of the procedures and the technical characteristics of the data produced seem ideal for use in routine, systematic formative evaluation. The functional utility of the measurement procedures in improving services remains to be determined.

References

Arter, J. A., & Jenkins, J. R. Differential diagnosis-prescriptive teaching: A critical appraisal (Technical Report No. 80). Urbana-Champaign, Ill.: University of Illinois, Center for the Study of Reading, 1978.

Bohannon, R. Direct and daily measurement procedures in the identification and treatment of reading behaviors of children in special education. Doctoral dissertation, University of Washington, 1975.

Brigham, T., Graubard, P., & Stans, A. Analysis of the effects of sequential reinforcement contingencies on aspects of composition. Journal of Applied Behavioral Analysis, 1972, 5, 421-429.

Carroll, J. B., Davies, P., & Richman, B. Word frequency book. Boston: Houghton Mifflin, 1971.

Crutcher, C. E., & Hofmeister, A. M. Effective use of objectives and monitoring. Teaching Exceptional Children, 1975, 7(2), 78-80.

Dale, E., & Chall, J. A. A formula for predicting readability. Educational Research Bulletin (Ohio State University), 27 (Jan. 21 and Feb. 17, 1948), 11-20, 37-54.

Deno, S., & Mirkin, P. Data based IEP development: An approach to substantive compliance. Teaching Exceptional Children, 1980, 12(3) 92-98.

Diederich, P. B. Problems and possibilities of research in the teaching of written composition. In Research design and the teaching of English. Champaign, Ill.: National Council of Teachers of English, 1964. (Eric Document Reproduction Service No. ED 001 917)

Dixon, E. A. Syntactic indexes and student writing performance. Elementary English, May 1972, 714-716.

Dunn, L. M., & Smith, J. O. Peabody language development kit. Minneapolis: American Guidance Service, 1967.

Durost, W. N., Bixler, H. H., Wrightstone, J. W., Prescott, G. A., & Balow, I. H. Metropolitan achievement test. New York: Harcourt Brace Jovanovich, 1971.

Federal Register. Department of Health, Education and Welfare, Washington, D.C., 42(163), Tuesday, August 23, 1977, Part II.

Finn, P. Computer-aided description of mature word choices in writing. In C. R. Cooper & L. Odell (Eds.), Evaluating writing: Describing, measuring, judging. Urbana, Ill.: National Council of Teachers of English, 1977.

Frumess, S. A comparison of management groups involving the use of the standard behavior chart. Unpublished doctoral dissertation, 1973.

Golub, L. S., & Kidder, C. L. Computer application of a syntactic density measure. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, April 1974.

Hammill, D. D., & Larsen, S. C. The test of written language. Austin, Tex., Pro-Ed, 1978.

Hillerich, R. L. Evaluation of written language. Elementary English, November 1971, 839.

Hunt, K. W. Grammatical structures written at three grade levels (Research Report No. 3). Champaign, Ill.: National Council of Teachers of English, 1965.

Hunt, K. W. Recent measures in syntactic development. Elementary English, XLII (November 1966), 732-739.

Kuehnle, K., Tindal, G., & Marston, D. Measures of children's writing. Unpublished research, 1979.

Lee, L., & Canter, S. M. Developmental sentence scoring. Journal of Speech and Hearing Disorders, 1971, 36, 335-340.

Lovitt, T., Schaff, M., & Sayre, E. The use of direct and continuous measurement to evaluate reading materials and procedures. Focus on Exceptional Children, 1970, 2, 1-11.

Madden, R., Gardner, E. F., Rudman, H. C., Karlsen, B., & Merwin, J. C. Stanford achievement test. New York: Harcourt Brace Jovanovich. 1978.

Mirkin, P., & Deno, S. Formative evaluation in the classroom: An approach to improving instruction (Research Report No. 10). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities, 1979.

Mullis, I. V. S. The primary trait system for scoring writing tasks. Denver, Colo.: Education Commission of the States, National Assessment of Educational Progress, 1976.

Myklebust, H. R. Development and disorders of written language. New York: Grune & Stratton, 1965.

Nystrand, M. Assessing written communicative competence: A textual cognition model. Report prepared at the Ontario Institute for Studies in Education, 1977.

O'Donnell, R. C., Griffin, W. T., & Norris, R. C. A transformational analysis of oral and written grammatical structures in the language of children in grades three, five, and seven. Journal of Educational Research, 1967, 61, 35-39.

Page, E. B. Analyzing student essays by computer. Internal Review of Education, 1968, 14, 216-217.

Perron, J. D. Written syntactic complexity and the modes of discourse. Paper presented at the Annual Meeting of the American Educational Research Association, New York, April 1977.

Rubin, R., & Buium, N. Language parameters in written compositions of nine-year old children. Paper presented at the annual meeting of the American Psychological Association, New Orleans, August 1974.

Rubin, R., Buium, N., & Balow, B. Grammatical sophistication in written compositions of nine-year-old children. National Institute of Education (DHEW), Interim Report 18, Project 6-1176, 1975.

Slotnick, L. B. Toward a theory of computer essay grading. Journal of Educational Measurement, 1972, 9(4), 253.

Stahl, A. Structural analysis of children's compositions. Research in Teaching English, 1974, 8, 184-205.

Thorpe, L. P., Lefever, D. W., & Naslund, A. SRA achievement series: Language arts. Chicago: Science Research Associates, 1968.

Veal, R. L. Syntactic measures and rated quality in the writing of young children (Report No. 8). University of Georgia, Dept. of Language Education, 1974.

White, O. R., & Haring, N. G. Exceptional teaching: A multimedia training package. Columbus, Ohio: Charles E. Merrill, 1976.

Table 1

Correlations Between Scores from a Writing Sample Using a Picture Stimulus and Scores on the TOWL*

TOWL Subtests	Mean T-unit Length	Mature Words	Large Words	Words Spelled Correctly	Total Words Written	Rate of Words Written
Vocabulary	.09[a]	.56	.42[b]	.68	.65	.27[a]
Thematic Maturity	.33[b]	.69	.29	.68[c]	.66	.19[a]
Spelling	.06[b]	.79	.61	.86	.75	.40[b]
Word Usage	.13[a]	.68	.47[c]	.77	.71	.35[b]
Style	.07[a]	.72	.53[c]	.74	.66	.33[b]
Raw Total	.13[a]	.78	.56	.88	.82	.36[b]
WLQ	.03[a]	.67	.56	.83	.77	.43[c]
Median Correlations	.09	.69	.53	.77	.71	.35

* N = 28. All correlations are significant at .001 unless otherwise noted.

[a]Not significant.

[b]Significant at .05.

[c]Significant at .01.

Table 2

Correlations Between Scores from a Writing Sample Using a Story Starter and Criterion Scores on the TOWL*

TOWL Subtests	Mean T-unit Length	Mature Words	Large Words	Words Spelled Correctly	Total Words Written
Vocabulary	.11[a]	.41[b]	.45[c]	.45[c]	.41[b]
Thematic Maturity	.26[a]	.65	.55[c]	.60	.59
Spelling	.25[a]	.71	.72	.73	.70
Word Usage	.18[a]	.57	.54	.57	.55
Style	.21[a]	.66	.65	.67	.63
Raw Total	.20[a]	.65	.66	.67	.63
WLQ	.22[a]	.67	.67	.64	.62
Median Correlations	.20	.65	.65	.64	.62

* N = 28. All correlations are significant at .001 unless otherwise noted.

[a]Not significant.

[b]Significant at .05.

[c]Significant at .01.

Table 3

Correlations Between Scores from a Writing Sample Using a Story Starter and Scores on Achievement Tests *

Dependent Measure	Sample Duration (Minutes) 1	2	3	4	5	Criterion Measure
Mean T-unit Length	.58	.42[b]	.34[b]	.41[b]	.36[b]	TOWL
	.52[c]	.39[b]	.22[a]	.24[a]	.32[b]	Stanford
Mature Words	.79	.77	.84	.85	.82	TOWL
	.69	.68	.71	.71	.72	Stanford
Large Words	.50[c]	.63	.65	.70	.70	TOWL
	.42[b]	.57	.59	.67	.68	Stanford
Large Words Modified	.45[c]	.58	.70	.70	.71	TOWL
	.44[c]	.63	.75	.76	.75	Stanford
Words Spelled Correctly	.74	.78	.79	.80	.78	TOWL
	.68	.69	.69	.68	.67	Stanford
Total Words Written	.73	.78	.81	.81	.79	TOWL
	.62	.68	.65	.63	.62	Stanford
Mean Correlations[d]	.60	.63	.65	.67	.67	

* All correlations are significant at .001 unless otherwise noted.

[a] Not significant.

[b] Significant at .05.

[c] Significant at .01.

[d] Mean correlations do not include T-unit data.

Table 4

Correlations Between Scores from a Writing Sample Using a Topic Sentence and Scores on Achievement Tests

Dependent Measure	Sample Duration (Minutes) 1	2	3	4	5	Criterion Measure
Mean T-unit Length	.30[a]	.19[a]	.37[b]	.16[a]	.16[a]	TOWL
	.15[a]	.03[a]	.20[a]	.09[a]	.07[a]	Stanford
Mature Words	.73	.73	.77	.77	.76	TOWL
	.61	.63	.68	.65	.60	Stanford
Large Words	.61	.61	.70	.74	.73	TOWL
	.42[c]	.44[c]	.56	.57	.55	Stanford
Large Words Modified	.59	.57	.62	.66	.69	TOWL
	.40[c]	.42[c]	.47[c]	.49[c]	.51[c]	Stanford
Words Spelled Correctly	.67	.72	.71	.71	.73	TOWL
	.64	.62	.60	.61	.63	Stanford
Total Words Written	.66	.68	.69	.68	.72	TOWL
	.62	.57	.57	.58	.61	Stanford
Mean correlations[d]	.60	.60	.64	.65	.65	

* All correlations are significant at .001 unle otherwise noted.

[a] Not significant.

[b] Significant at .05.

[c] Significant at .01.

[d] Mean correlations do not include T-unit data.

Table 5

Correlations Between Scores from a Writing Sample Using a Picture Stimulus and Scores on Achievement Tests

Dependent Measure	Sample Duration (Minutes) 1	2	3	4	5	Criterion Measure
Mean T-unit Length	.60	.49[c]	.47[c]	.29[a]	.18[a]	TOWL
	.45[c]	.49[c]	.35[b]	.21[a]	.02[a]	Stanford
Mature Words	.65	.77	.86	.88	.83	TOWL
	.52[c]	.67	.74	.77	.73	Stanford
Large Words	.50[c]	.68	.70	.70	.75	TOWL
	.50[c]	.60	.68	.68	.72	Stanford
Large Words Modified	.39[b]	.60	.68	.68	.70	TOWL
	.38[b]	.52[c]	.60	.63	.64	Stanford
Words Spelled Correctly	.68	.71	.78	.78	.80	TOWL
	.65	.67	.74	.77	.76	Stanford
Total Words Written	.63	.72	.74	.75	.75	TOWL
	.56	.67	.67	.71	.70	Stanford
Mean correlations	.55	.66	.72	.74	.73	

* All correlations are significant at .001 unless otherwise noted.

[a] Not significant.

[b] Significant at .05.

[c] Significant at .01.

Table 6

Correlations Between Various Writing Stimulus Formats

on Four Behavioral Measures

Dependent Measure	Story Starter and Topic Sentence	Story Starter and Picture Stimulus	Topic Sentence and Picture Stimulus
Mature Words	.75	.79	.74
Large Words	.59	.55	.50
Words Spelled Correctly	.81	.87	.86
Total Words Written	.79	.86	.85

Table 7

Mean Performance of 28 Elementary Students on Behavioral Measures of Written Expression and Elicited by Story Starters, Topic Sentences, and Picture Stimuli

Dependent Measure	Story Starter	Topic Sentence	Picture Stimulus
Mature Words	10.8	9.7	9.8
Words Spelled Correctly	40.5	40.0	37.4
Total Words Written	44.8	42.5	41.1

Table 8

Correlations Between Scores from a Writing Sample Using a Story Starter and Developmental Sentence Scoring Data for Various Sample Durations

Dependent Measure	Sample Duration (Minutes) 1	2	3	4	5	Criterion Measure
Mean T-unit Length					.29[b]	DSS
Mature Words	.54	.66	.72	.73	.74	DSS
Large Words	.38	.44	.48	.47	.47	DSS
Large Words Modified	.23[a]	.31[b]	.35	.35	.34	DSS
Words Spelled Correctly	.67	.81	.84	.87	.76	DSS
Total Words	.65	.80	.84	.88	.84	DSS
Letters in Correct Sequence	.64	.76	.78	.83	.86	DSS
Average Correlation[c]	.52	.63	.67	.69	.67	

*All correlations are significant at .001 unless otherwise indicated.

[a]Significant at .05.

[b]Significant at .01.

[c]Average correlations do not include T-unit data.

Table 9

Correlations Between Scores from a Writing Sample Using a Story Starter and Scores from the TOWL for Combined Samples*

TOWL Subtest	Mature Words	Large Words	Words Spelled Correctly	Total Words Written	Letters in Correct Sequence
Vocabulary	.56	.43	.56	.56	.63
Thematic Maturity	.70	.56	.60	.62	.65
Spelling	.75	.67	.68	.68	.70
Word Usage	.74	.63	.69	.70	.74
Style	.63	.53	.59	.61	.57
Total	.76	.63	.71	.72	.75

[a]N = 135. All correlations are significant at .001.

Table 10

Performance Means of Students from Regular (N=44) and Resource Classrooms (N=90) on Dependent Measures

Measure	Group	Mean	Standard Deviation	T-Value	Probability
Total Written	Resource	31.9	19.8	-5.64	.001
	Regular	53.3	21.3		
Mature Words	Resource	6.3	5.1	-5.52	.001
	Regular	12.4	6.4		
Words Spelled Correctly	Resource	25.8	19.3	-6.05	.001
	Regular	49.1	21.9		
Letters in Sequence	Resource	130.9	97.0	-4.68	.001
	Regular	227.6	95.3		

Table 11

Correlation Coefficients for LD and Non-LD Groups

on Scores from Written Samples and the TOWL*

Group	Mature Words	Words Spelled Correctly	Total Words Written	Letters in Correct Sequence	TOWL Subtest
LD	.56	.50	.60	.70	Vocabulary
Non-LD	.43	.44	.38	.46	
LD	.64	.50	.58	.68	Thematic Maturity
Non-LD	.63	.47	.48	.53	
LD	.73	.66	.69	.79	Spelling
Non-LD	.67	.54	.55	.57	
LD	.72	.63	.73	.84	Word Usage
Non-LD	.64	.56	.56	.60	
LD	.60	.62	.62	.62	Style
Non-LD	.52	.43	.48	.43	
LD	.77	.67	.75	.83	Subtest Total
Non-LD	.72	.60	.61	.63	

* All correlations significant at .001 level.

Table 12

Mean Grade Level Performance and Standard Deviations for Combined Sample on Four Behavioral Measures of Written Expression

	N	Mean	Standard Deviation
Total Words Written			
Grade 3	20	26.1	20.1
Grade 4	46	47.6	20.2
Grade 5	33	49.9	22.2
Grade 6	31	54.1	22.5
Mature Words			
Grade 3	20	6.1	5.3
Grade 4	46	10.4	6.7
Grade 5	33	11.2	6.2
Grade 6	31	12.3	6.4
Words Spelled Correctly			
Grade 3	20	22.9	19.8
Grade 4	46	41.6	20.5
Grade 5	33	45.3	23.5
Grade 6	31	49.1	24.9
Letters Written in Correct Sequence			
Grade 3	15	99.5	90.3
Grade 4	32	198.1	93.1
Grade 5	25	242.0	101.9
Grade 6	25	227.9	87.9

Table 13

Means and Standard Deviations for Resource and Regular Students on Total Words Written

	Resource Students			Regular Students		
	N	Mean	Standard Deviation	N	Mean	Standard Deviation
Grade 3	9	9.0	9.1	11	40.1	15.0
Grade 4	10	34.5	13.5	36	51.1	20.4
Grade 5	11	39.9	21.6	22	54.9	21.2
Grade 6	14	41.6	14.4	17	64.4	23.1

Table 14

Means and Standard Deviations for Resource and Regular Students on Mature Words Written

	Resource Students			Regular Students		
	N	Mean	Standard Deviation	N	Mean	Standard Deviation
Grade 3	9	1.7	2.7	11	9.6	4.1
Grade 4	10	5.1	3.6	36	11.9	6.7
Grade 5	11	8.5	5.9	22	12.5	6.0
Grade 6	14	9.4	4.0	17	14.6	7.0

Table 15

Means and Standardized Deviations for Resource and Regular Students on Words Spelled Correctly

	Resource Students			Regular Students		
	N	Mean	Standard Deviation	N	Mean	Standard Deviation
Grade 3	9	5.3	7.2	11	36.8	15.1
Grade 4	10	28.9	15.2	36	45.1	20.6
Grade 5	11	32.0	21.0	22	52.0	22.2
Grade 6	14	34.0	17.5	17	61.5	23.5

Table 16

Means and Standard Deviations for Resource and Regular Students on Letters Written in Correct Sequence

	Resource Students			Regular Students		
	N	Mean	Standard Deviation	N	Mean	Standard Deviation
Grade 3	6	18.3	29.7	9	153.5	74.0
Grade 4	5	113.4	72.9	27	213.9	89.2
Grade 5	8	208.0	101.3	17	258.1	101.1
Grade 6	10	164.8	55.2	15	270.0	80.9

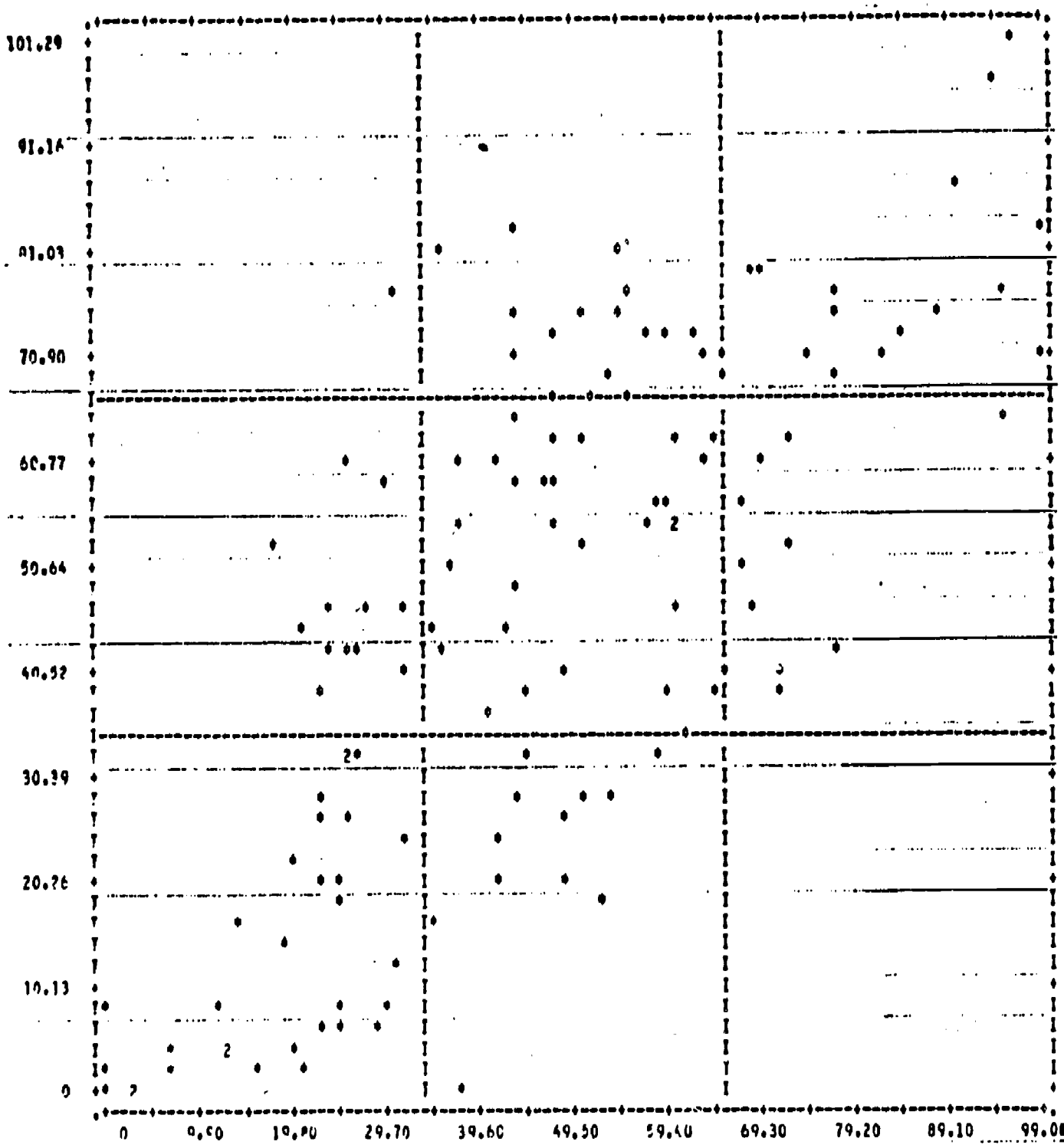


Figure 1 Scattergram illustrating the relationship between Total Words Written and Subtest Total from TOWL for 135 subjects.

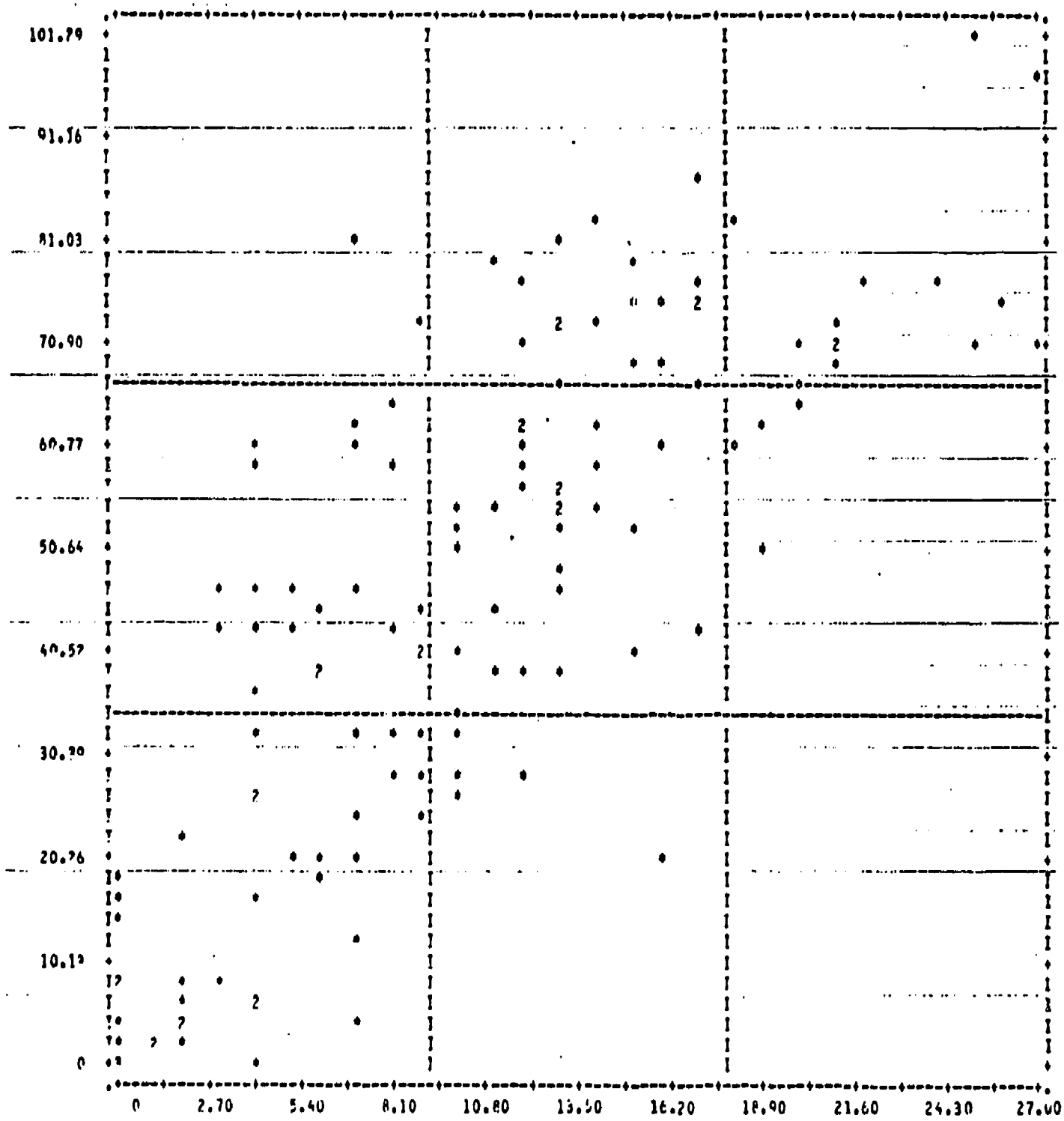
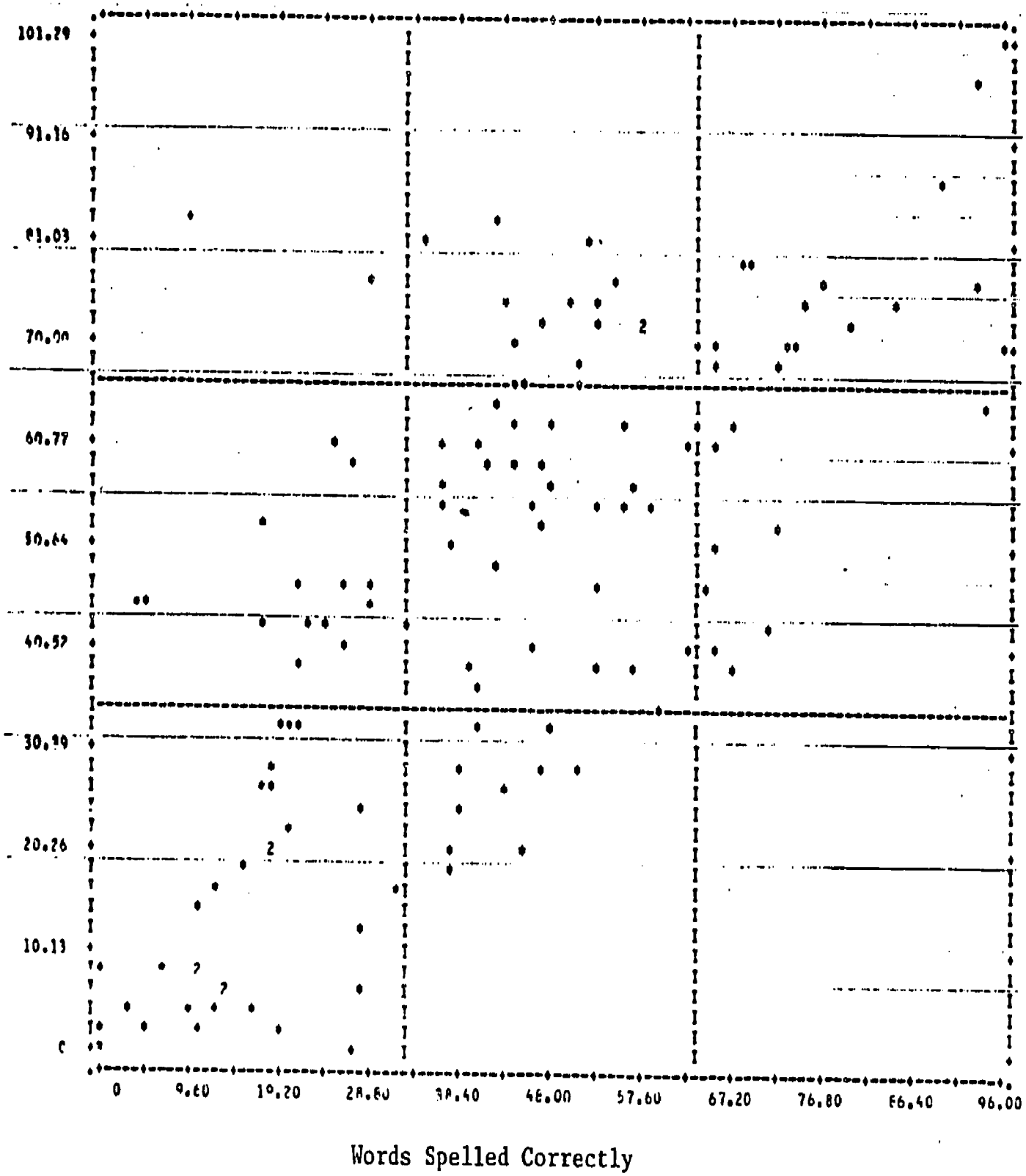


Figure 2 Scattergram illustrating the relationship between Mature Words Written and Subtest Total from TOWL for 135 subjects.



ubtest
otal
rom
OWL

Figure 3 Scattergram illustrating the relationship between Words Spelled Correctly and Subtest Total from TOWL for 135 subjects.

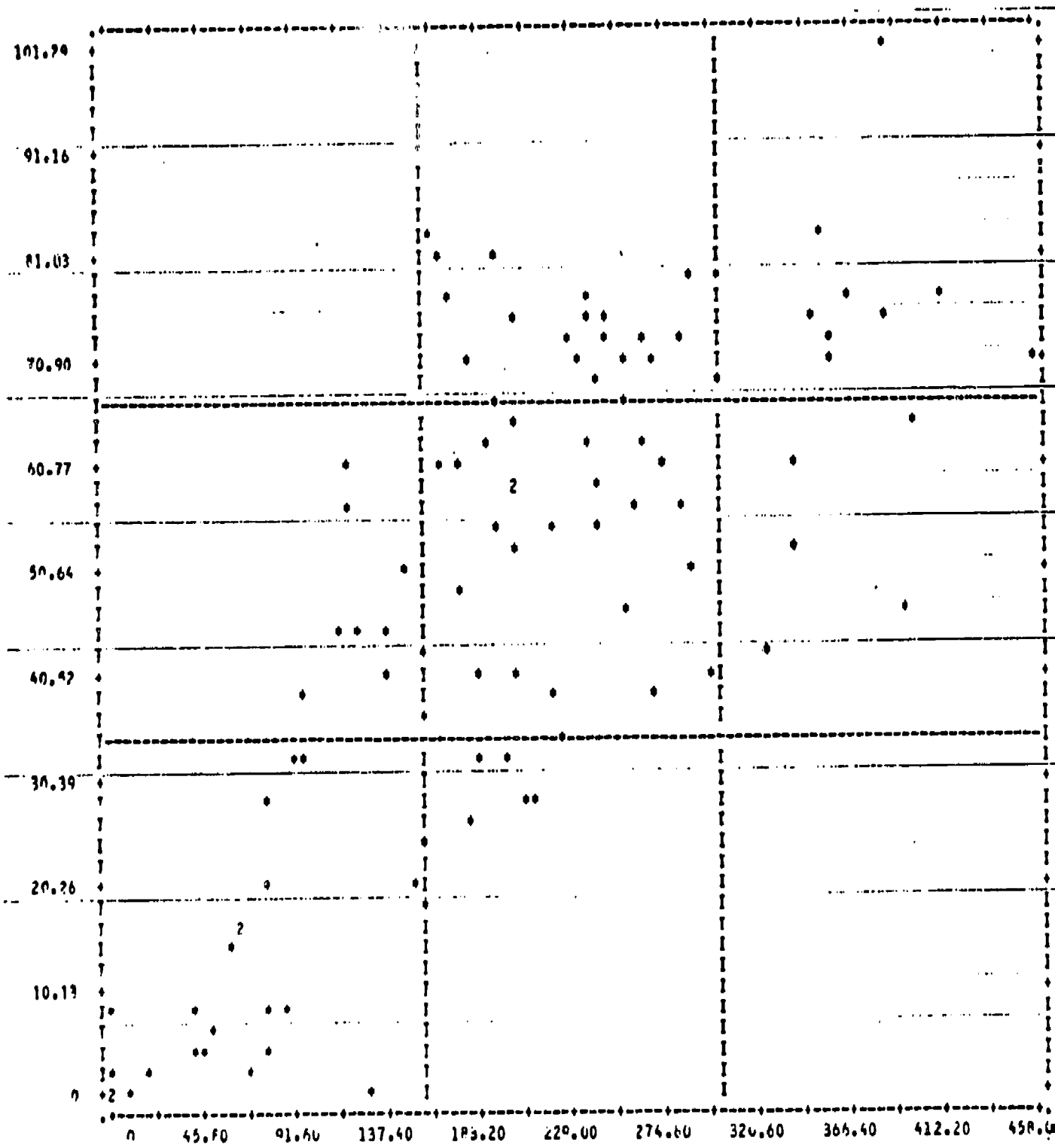


Figure 4 Scattergram illustrating the relationship between Letters Written in Correct Sequence and Subtest Total from TOWL for 135 subjects.

APPENDIX A

Examples of Story Starters and Topic Sentences

Story Starters

1. Tell a story about the night you were camping in the woods and you heard strange noises.
2. Tell a story about what kind of car you would buy.
3. Pretend that you are stranded on a tropical island by yourself. Tell a story about what happens to you.

Topic Sentences

1. Describe your favorite season of the year and tell why.
2. Write about summer vacation.
3. Write about what makes you happy.

APPENDIX B

Example of How T-Unit Length is Computed for a Written Composition

Example of T-Unit Scoring

[When I went on a trip to Texas I rode on an airplane.] [I had a really good time.] [I stayed in a hotel for three days] [and I went to a lot of parks.]

T-units = 4

Words = 34

Mean T-unit Length = 8.5

APPENDIX C

Finn's Undistinguished Word List and An Example of How a Written Composition is Scored for Mature Words

Undistinguished Word Choices

a
able
about
again
air
all
also
and
animals
another
any
anything
are
around
as
ask
asked
at
away
bad
be
because
been
before
being
better
big
build
busi..ess
but
buy
by
call
called
came
can
children
cleaning
clear
close
closed
come
could
couldn't
day
did
didn't
do
doesn't
doing
done
don't
down
drink
dumping
each
earth
either
else
enough
even
every
everyone
everything
fair
family
feel
few
filter
find
first
fish
fishing
food
for
form
from
get
getting
give
go
going
good
got
had
happen
happy
hard
have
having
he
health
help
him
his/her
home
house
how
I
if
I'm
important
in
into
is
it
its
job
just
keep
kill
know
lake
lakes
let
like
little
live
living
long
look
lot
make
making
man
many
may
maybe
me
mean
men
might
money
months
more
most
much
must
my
near
need
never
new
next
no
not
now
of
off
on
once
one hundred
open
or
other
our
out
over
own
part
pay
people
person
place
plants
problem
put
rather
reason
right
river
run
running
said
same
save
say
see
should
since
so
some
someone
something
soon
source
start
stay
still
stop
support
swim
take
talking
that
that's
than
the
their
them
then
there
they
thing
things
think
this
tune
to
too
try
until
up
us
use
very
want
was
water
way
we
well
what
when
where
who
why
will
with
without
won't
work
worked
working
world
would
years
you
your

Example of Mature Word Scoring

When I went on a trip to Texas I rode on an airplane. I had a really good time. I stayed in a hotel for 3 days and I went to a lot of parks.

Mature Words = 8

APPENDIX D

Example of How a Written Composition is Scored for Total Words Written

Example of Total Words Written Scored

When I went on a trip to Texas I rode on an airplane. I had a really good time. I stayed in a hotel for 3 days and I went to a lot of parks.

Total Words Written = 33

APPENDIX E

Example of How Large Words are Scored

Example of Large Words Scores

When I went on a trip to Texas I rode on an airplane. I had a really good time. I stayed in a hotel for three days and I went to a lot of parks.

Large Words = 1

APPENDIX F

Example of How Words Spelled Correctly are Scored

When I went on a trip to Teksas I road on an aerplane. I had a reely good time. I staid in a hotel for 3 days and I went to a lot of parks.

Words Spelled Correctly = 28

APPENDIX G

Scoring Categories Used in the Developmental Sentence Scoring System (Lee & Canter, 1971)

Category	Score	Criteria
Indefinite Pronouns or Noun Modifiers	1	it, this, that
	2	no, some, more, all, etc.
	3	something, somebody, someone
	4	nothing, nobody, no one, none
	5	any, anything, anybody, anyone, every, etc.
	6	both, few, many, each, several, most, etc.
Personal Pronouns	1	1st and 2nd person (I, me, you, etc.)
	2	3rd person (he, him, his, she, etc.)
	3	plural pronouns (we, us, they, etc.)
	4	those, these
	5	reflexive pronouns (myself, yourself, etc.)
	6	wh-pronouns (who, which, etc.); wh-word + infinitive
	7	(his) own, one, oneself, whichever, etc.
Main Verbs	1	uninflected verb; copula (is, 's)
	2	is + verb + ing
	3	-s, -ed; irregular past; copula am, are, was, were; auxiliary am, are, was, were
	4	can, will, may + verb; obligatory do + verb; emphatic do + verb
	5	could, would, should, might + verb; obligatory does, did + verb; emphatic does, did + verb
	6	must, shall + verb; have + verb + en; have ('ve) got
	7	passive, any tense
	8	have (had) been + verb + ing; modal + have + verb + en; modal + be + verb + ing; other auxiliary combinations

Category	Score	Criteria
Secondary Verbs	1	early developing infinitive complements
	2	noncomplementing infinitives
	3	participle, present or past
	4	early infinitival complements with differing subjects in kernals; later infinitival complements; obligatory deletions; infinitive with wh-word
	5	passive infinitival complement
	6	gerund
Negatives	1	it, this, that + copula or auxiliary is, 's + not
	2	can't, don't
	3	isn't, won't
	4	copula - negative or auxiliary-negative contractions; pronoun-auxiliary contraction + not; uncontracted negatives
	5	negatives with have; auxiliary have-negative contraction; pronoun-auxiliary have contraction
Conjunction	1	and
	2	but
	3	because
	4	so, and so, so that, if
	5	or, except, only
	6	where, when, for, till, since, as, etc. + adjective + as, as if, etc.; obligatory deletions; optional deletions; wh-words + infinitive
	7	therefore, however, whenever, etc.

Category	Score	Criteria
Interrogative Reversals	1	reversal of copula
	2	reversal of auxiliary be
	3	obligatory do, does, did; reversal of modal; tag question
	4	reversal of auxiliary have; reversal with any two auxiliaries
	5	reversal of three auxiliaries
Wh-Questions	1	who, what, what + noun
	2	where, how many, how much, etc.
	3	when, how, how + adjective
	4	why, what if, how come, how about + gerund
	5	whose, which, which + noun

APPENDIX H

Copy of the Story Starter Used in Study I

Story Starter Form B

Date _____
ID # _____

Pretend that you can travel anywhere that you want. Where would you go? How would you get there and what would you do when you got there?

APPENDIX I

Copies of the Story Starters

and Topic Sentences Used in Study II

Form B-1 Story Starter

Name _____

Tell a story about the night you were camping in the woods and you heard strange noises in the woods.

TW	MWC1	MWC2	BW1	BW2	TU1	TU2	SP1	TIME

Form B-2 Story Starter

Name _____

Tell a story about what kind of car you would buy and what features it would have.

TW	MWC1	MWC2	BW1	BW2	TU1	TU2	SP1	TIME

Form B-3 Story Starter

Name _____

Pretend that you are stranded on a tropical island by yourself.

Tell a story about what happens to you.

TW	MWC1	MWC2	BW1	BW2	TU1	TU2	SP1	TIME

Form B-1 Topic Sentence Name _____

Describe your favorite season of the year and tell why.

TW	MWC1	MWC2	BW1	BW2	TU1	TU2	SP1	TIME

Form B-2 Topic Sentence Name _____

Imagine that you could travel anywhere in the world that you wanted.
Where would you go? Why?

TW	MWC1	MWC2	BW1	BW2	TU1	TU2	SP1	TIME

Form B-3 Topic Sentence

Name _____

What is your favorite Holiday of the year? Give your reasons.

TW	MWC1	MWC2	BW1	BW2	TU1	TU2	SP1	TIME

APPENDIX J

Order of Presentation of Stimulus Materials

Order for 1/2 of subject sample

1. Story Starter
2. Stanford Subtest
3. Topic Sentence
4. Picture Stimulus
5. Test of Written Language

Order for other 1/2 of subject sample

1. Topic sentence
2. Stanford Subtest
3. Story Starter
4. Picture Stimulus
5. Test of Written Language

APPENDIX K

Copies of the Story Starters and Topic Sentences Used in Study III and an Illustration of How a Written Composition is Scored for Letters in Correct Sequence

Write a story that begins with: One summer I went on a trip.

TW	MWC1	MWC2	BW1	BW2	TU1	TU2	SP1	TIME

Form B-2 Story Starter

Name _____

Write a story that begins with: One night I went outside when it was very dark.

TW	MWC1	MWC2	BW1	BW2	TU1	TU2	SP1	TIME

Form B-3 Story Starter Name _____

Write a story that begins with: One day something happened which made me very happy.

TW	MWC1	MWC2	BW2	TU1	TU2	SP1	TIME

Form B-1 Topic Sentence Name _____

Write about ....."Summer vacation."

TW	MWC1	MWC2	BW1	BW2	TU1	TU2	SP1	TIME

Form B-2 Topic Sentence Name _____

Write about .....your neighborhood, where you live.

TW	MWC1	MWC2	BW1	BW2	TU1	TU2	SP1	TIME

Form B-3 Topic Sentence

Name _____

Write about what makes you happy.

TW	MWC1	MWC2	BW1	BW2	TU1	TU2	SP1	TIME

5 2 5 3 2 5 2 4 2 2
^WHEN^ I WENT ON A TRIP TO ^TEKSAS^ I ^ROAD

3 3 7 2 4 2 4 5 5
ON AN ^AERPLANE. I HAD A ^REELY GOOD TIME.

2 4 3 2 6 4 5 4 2
I ^STAID^ IN A HOTEL FOR 3 DAYS AND I

5 3 2 4 3 6
WENT TO A LOT OF PARKS.

LETTERS IN CORRECT SEQUENCE = 122

APPENDIX L

Analysis of Variance Tables for

One-way and Two-way ANOVAs

Table A

Oneway ANOVA of Grades 3, 4, 5, and 6 for Dependent Measure,

Total Words Written.

	Sum of Squares	df	Mean Squares	F	Significance
Main Effect due to Grade Level	9676.9	3	3225.6	6.9	.001
Residual	44418.7	95	467.6		
Total	54095.7	98			

Table B

Oneway ANOVA of Grades 3, 4, 5, and 6 for Dependent Measure,

Mature Words Written

	Sum of Squares	df	Mean Squares	F	Significance
Main Effect due to Grade Level	416.6	3	138.9	3.5	.019
Residual	3771.0	95	39.7		
Total	4187.6	98			

Table C

Oneway ANOVA of Grades 3, 4, 5, and 6 for Dependent Measure,

Words Spelled Correctly

	Sum of Squares	df	Mean Squares	F	Significance
Main Effect due to Grade Level	8116.7	3	2705.5	5.2	.002
Residual	49397.9	95	519.9		
Total	57514.6	98			

Table D

Oneway ANOVA of Grades 3, 4, 5, and 6 for Dependent Measure

Letters Written in Correct Sequence

	Sum of Squares	df	Mean Squares	F	Significance
Main Effects due to Grade Level	4938.5	3	1646.2	3.5	.018
Residual	44607.5	95	469.5		
Total	49546.0				

Table E

Two Way ANOVA Effect of Dependent Variable

Total Words Written on Grade Level (3, 4, 5, or 6)

and Educational Status (Regular or Resource Student)

	Sum of Squares	df	Mean Squares	F	Significance
Main Effects					
Grade Level	10995.3	3	3665.1	10.4	.001
Educational Status	8121.3	1	8121.3	23.1	.001
Interaction	1140.7	3	380.2	1.0	.359
Residual	30821.9	88	350.2		
Total	50863.7	95			

Table F

Two Way ANOVA Effect of Dependent Variable Mature Words Written on Grade Level (3, 4, 5, or 6) and Educational Status (Regular or Resource Student)

	Sum of Squares	df	Mean Squares	F	Significance
Main Effects					
Grade Level	520.5	3	173.5	5.6	.001
Educational Status	650.5	1	650.5	21.0	.001
Interaction	58.9	3	19.6	.63	.594
Residual	2724.0	88			
Total	3925.9	95			

Table G

Two Way ANOVA Effect of Dependent Variable Words Spelled on Grade Level (3, 4, 5, or 6) and Educational Status (Regular or Resource Student)

	Sum of Squares	df	Mean Squares	F	Significance
Main Effects					
Grade Level	10347.9	3	3449.3	9.2	.001
Educational Status	11317.2	1	11317.2	30.3	.001
Interaction	810.7	3	270.2	.726	.539
Residual	32762.4	88			
Total	54373.4	95			

Table H

Two Way ANOVA Effect of Dependent Variable Letters Written in Correct Sequence on Grade Level (3, 4, 5, or 6) and Educational Status (Regular or Resource Student)

	Sum of Squares	df	Mean Squares	F	Significance
Main Effects					
Grade Level	214040.3	3	7.346.7	9.9	.001
Educational Status	170089.5	1	170089.5	23.7	.001
Interaction	17615.7	3	5871.9	.819	.487
Residual	400584.1	88			
Total	630892.5	95			

PUBLICATIONS

Institute for Research on Learning Disabilities
University of Minnesota

The Institute is not funded for the distribution of its publications. Publications may be obtained for $3.00 per document, a fee designed to cover printing and postage costs. Only checks and money orders payable to the University of Minnesota can be accepted. All orders must be prepaid.

Requests should be directed to: Editor
IRLD
350 Elliott Hall
75 East River Road
University of Minnesota
Minneapolis, Minnesota 55455

Ysseldyke, J. E. Assessing the learning disabled youngster: The state of the art (Research Report No. 1). November, 1977.

Ysseldyke, J. E., & Regan, R. R. Nondiscriminatory assessment and decision making (Monograph No. 7). February, 1979.

Foster, G., Algozzine, B., & Ysseldyke, J. Susceptibility to stereotypic bias (Research Report No. 3). March, 1979.

Algozzine, B. An analysis of the disturbingness and acceptability of behaviors as a function of diagnostic label (Research Report No. 4). March, 1979.

Algozzine, B., & McGraw, K. Diagnostic testing in mathematics: An extension of the PIAT? (Research Report No. 5). March, 1979.

Deno, S. L. A direct observation approach to measuring classroom behavior: Procedures and application (Research Report No. 6). April, 1979.

Ysseldyke, J. E., & Mirkin, P. K. Proceedings of the Minnesota roundtable conference on assessment of learning disabled children (Monograph No. 8). April, 1979.

Somwaru, J. P. A new approach to the assessment of learning disabilities (Monograph No. 9). April, 1979.

Algozzine, B., Forgnone, C., Mercer, C. D., & Trifiletti, J. J. Toward defining discrepancies for specific learning disabilities: An analysis and alternatives (Research Report No. 7). June, 1979.

Note: Monographs No. 1 - 6 and Research Report No. 2 are not available for distribution. These documents were part of the Institute's 1979-1980 continuation proposal, and/or are out of print.

Algozzine, B. The disturbing child: A validation report (Research Report No. 8). June, 1979.

Ysseldyke, J. E., Algozzine, B., Regan, R., & Potter, M. Technical adequacy of tests used by professionals in simulated decision making (Research Report No. 9). July, 1979.

Jenkins, J. R., Deno, S. L., & Mirkin, P. K. Measuring pupil progress toward the least restrictive environment (Monograph No. 10). August, 1979.

Mirkin, P. K., & Deno, S. L. Formative evaluation in the classroom: An approach to improving instruction (Research Report No. 10). August, 1979.

Thurlow, M. L., & Ysseldyke, J. E. Current assessment and decision-making practices in model programs for the learning disabled (Research Report No. 11). August, 1979.

Deno, S. L., Chiang, B., Tindal, G., & Blackburn, M. Experimental analysis of program components: An approach to research in CSDC's (Research Report No. 12). August, 1979.

Ysseldyke, J. E., Algozzine, B., Shinn, M., & McGue, M. Similarities and differences between underachievers and students labeled learning disabled: Identical twins with different mothers (Research Report No. 13). September, 1979.

Ysseldyke, J., & Algozzine, R. Perspectives on assessment of learning disabled students (Monograph No. 11). October, 1979.

Poland, S. F., Ysseldyke, J. E., Thurlow, M. L., & Mirkin, P. K. Current assessment and decision-making practices in school settings as reported by directors of special education (Research Report No. 14). November, 1979.

McGue, M., Shinn, M., & Ysseldyke, J. Validity of the Woodcock-Johnson psycho-educational battery with learning disabled students (Research Report No. 15). November, 1979.

Deno, S., Mirkin, P., & Shinn, M. Behavioral perspectives on the assessment of learning disabled children (Monograph No. 12). November, 1979.

Sutherland, J. H., Algozzine, B., Ysseldyke, J. E., & Young, S. What can I say after I say LD? (Research Report No. 16). December, 1979.

Deno, S. L., & Mirkin, P. K. Data-based IEP development: An approach to substantive compliance (Monograph No. 13). December, 1979.

Ysseldyke, J., Algozzine, B., Regan, R., & McGue, M. The influence of test scores and naturally-occurring pupil characteristics on psycho-educational decision making with children (Research Report No. 17). December, 1979.

Algozzine, B., & Ysseldyke, J. E. Decision makers' prediction of students' academic difficulties as a function of referral information (Research Report No. 18). December, 1979.

Ysseldyke, J. E., & Algozzine, B. Diagnostic classification decisions as a function of referral information (Research Report No. 19). January, 1980.

Deno, S. L., Chiang, B., Mirkin, P. K., & Lowry, L. Relationships among simple measures of reading and performance on standardized achievement tests (Research Report No. 20). January, 1980.

Deno, S. L., Lowry, L., Mirkin, P. K., & Kuehnle, K. Relationships among simple measures of spelling and performance on standardized achievement tests (Research Report No. 21). January, 1980.

Deno, S. L., Marston, D., & Mirkin, P. K. Relationships among simple measures of written expression and performance on standardized achievement tests (Research Report No. 22). January, 1980.

Mirkin, P. K., Deno, S. L., Tindal, G., & Kuehnle, K. Formative evaluation: Continued development of data utilization systems (Research Report No. 23). January, 1980.

Deno, S. L., Robinson, S., Evans, P., & Mirkin, P. K. Relationships among classroom observations of social adjustment and sociometric rating scales (Research Report No. 24). January, 1980.

Thurlow, M. L., & Ysseldyke, J. E. Factors influential on the psychoeducational decisions reached by teams of educators (Research Report No. 25). February, 1980.

Ysseldyke, J. E., & Algozzine, B. Diagnostic decision making in individuals susceptible to biasing information presented in the referral case folder (Research Report No. 26). March, 1980.

Thurlow, M. L., & Greener, J. W. Preliminary evidence on information considered useful in instructional planning (Research Report No. 27). March, 1980.

Ysseldyke, J. E., Regan, R. R., & Schwartz, S. Z. The use of technically adequate tests in psychoeducational decision making (Research Report No. 28). April, 1980.

Richey, L., Potter, M., & Ysseldyke, J. Teachers' expectations for the siblings of learning disabled and non-learning disabled students: A pilot study (Research Report No. 29). May, 1980.