

THE UNIVERSITY OF ALBERTA

A STUDY OF THE RELATIONSHIP BETWEEN OBJECTIVE
TESTS AND WRITTEN ESSAYS AS MEASURES OF THE
WRITING ABILITY OF GRADE TEN STUDENTS

by

EDWARD AUSTIN JONES

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF SECONDARY EDUCATION

EDMONTON, ALBERTA

FALL, 1969

ABSTRACT

The study was an investigation of objective tests and essay writing as measures of writing ability. Three objectives were basic to the study:

1. To establish a reliable score of students' writing performance to serve as a criterion against which to validate five objective tests of writing.
2. To determine, both individually and in various combinations, the relative validities of five objective tests of writing.
3. To note the effect of the addition of a twenty-minute paragraph to the validities of the selected combinations of objective tests.

A group of 151 tenth-grade students wrote on three different essay topics and took five objective tests of writing skills. Each of four raters assigned scores from one to four to each of the three essays. The total of twelve scores thus assigned became the criterion for validating the five objective tests, first individually and then in various combinations. Later the total essay score on two essays became the criterion when the third essay (a twenty minute paragraph) was added as a predictor and the variables were again examined individually and in combinations. Of the five objective tests used, three were selected from available standardized tests (STEP Writing: Form 2B; Cooperative English Tests: Reading Comprehension, Form 2A; and Cooperative English Tests:

English Expression, Form 2A) and two were constructed by the investigator (Written Expressional Vocabulary, and Sentence Sensitivity). Finally, sex and intelligence were included as predictors and the unique contribution of each variable to the essay score variance was re-examined.

A three-way analysis of variance of the ratings assigned by four raters on the three essays revealed an estimated reading reliability of .89 and an estimated score reliability of .82. These reliabilities established the total score on all three essays as a relatively sound criterion for validating the five objective tests. The reliability of essay scores was shown to be primarily a function of the number of different essays and the number of different readings. The presence of a rater effect and a rater by essay interaction were also observed.

All five objective tests appeared to have acceptable validity, with coefficients ranging from .569 to .622 with the criterion. Composites of two different objective tests produced multiple correlations ranging from .639 to .702. The three combinations with the highest correlations contained the investigator's Sentence Sensitivity Test. Eight of the nine composites of three different objective tests produced correlations above .710. Finally, the combination of all five objective tests produced a multiple correlation of .758 with the three-essay criterion. As hypothesized both the test of Written Expressional Vocabulary and the

test of Sentence Sensitivity constructed by the investigator made significant contributions to the variance of students' composition scores when each was examined in the presence of the other four objective tests. The unique contributions of the English Expression Test and the STEP Writing Test were also significant.

When the paragraph was removed from the three-essay criterion and the criterion reduced to the total scores on two essays, it was found that composites of two objective tests produced validity coefficients ranging from .569 to .662. With the addition of the paragraph to the same ten two-test composites the correlations ranged from .665 to .721, with an average increase of .070 per correlation coefficient. Similarly, the addition of the paragraph to composites of three objective tests raised all correlation coefficients. The combination of all five tests plus the paragraph produced a multiple correlation of .739 with the two-essay criterion. As hypothesized, the twenty-minute paragraph contributed significantly to the prediction of composition scores when the relationship was measured in the presence of the five objective tests. The unique contributions of the Sentence Sensitivity Test and of the English Expression Test were also significant.

Finally, the investigation showed that the addition of the variables of sex and intelligence as predictors made no significant contribution to the prediction of composition scores on either the three- or two-essay criterion.

ACKNOWLEDGEMENTS

The writer wishes to acknowledge the help and encouragement accorded him during the conduct of the investigation. Grateful appreciation is especially extended to Dr. Earl W. Buxton, the chairman of my committee, whose scholarship, insight, and encouragement were invaluable.

Discussions with Dr. Robert D. Armstrong were most helpful and are acknowledged with thanks. The writer is indebted, also, to Dr. Patricia Lane McFetridge and Dr. Jean D. Crowther for their constructive criticisms of the study, and to Professor James B. Bell for his assistance when the study was initiated.

The writer is also indebted to Dr. Gerald Berry, Head of the Department of Secondary Education and to the University of Alberta for financial assistance during my studies.

Appreciation is also expressed to the Edmonton Public School Board and to the staff and students of the participating schools for their cooperation in the collection of the data.

A special thank you is given to my wife, Pauline, for her helpful suggestions and discussions during the planning and writing of the study.

TABLE OF CONTENTS

CHAPTER	PAGE
I. INTRODUCTION	1
The Debate: Objective vs. Essay Tests of Writing . . .	1
The General Problem	4
Purposes of the Study	5
Hypotheses	7
Definitions	9
Limitations	10
Overview of the Investigation	12
II. CONCEPT OF WRITING ABILITY AND THE WRITING PROCESS . . .	14
Introduction	14
Statements of Leading Educators	15
Writers on Writing	21
General Statement of Concept of Writing Ability and Writing Process	24
Informations from Related Disciplines	27
Semantics and Written Composition	27
Logic and Written Composition	29
Psychology, Psycholinguistics and Composition . . .	31
Rhetoric and Written Composition	35
Linguistics (Grammar) and Written Composition . . .	37
Further Explication of Concept of Writing	40
Summary	43

CHAPTER	PAGE
III. OBJECTIVE TESTS	45
Review: Objective Tests of Writing Ability	45
Summary	52
The Selection of Standardized Tests for this	
Investigation	53
Sequential Tests of Educational Progress: Writing,	
Form 2B.	53
Description	53
Validity	54
Reliability	55
Cooperative English Tests: Reading Comprehension,	
Form 2A	56
Description	56
Justification for Its Inclusion	57
Validity	62
Reliability	64
Cooperative English Tests: English Expression,	
Form 2A	64
Summary	66
The Construction of Objective Tests for this	
Investigation	67
Written Expressional Vocabulary	67
Justification for the Test	67

CHAPTER	PAGE
Description of the Test	73
Procedure for Construction	76
Sentence Sensitivity Test	77
Sentence Structure and Prose Quality	77
Description of the Test	84
Procedure for Construction	87
Pilot Study	87
IV. THE CRITERION MEASURE	91
Variables Associated with Essay Testing	91
The Assignment Variable	91
Procedure to Minimize the Assignment Variable . .	92
The Writer Variable	93
Procedure to Minimize the Writer Variable	95
The Rater Variable	96
Procedure to Minimize the Rater Variable	98
Holistic vs. Analytic Rating of Essays	98
Description of Holistic Method Used	102
Procedure Used for Grading	104
Qualifications of Raters	104
Summary of Procedures Followed in Establishing	
Criterion	105

CHAPTER	PAGE
V. DESIGN OF THE STUDY	107
Sources of Data	107
The Sample Tested	108
Test Administration	110
Treatment of Data	110
VI. RESULTS: THE CRITERION MEASURE	114
VII. RESULTS: RELATIONS BETWEEN PREDICTORS AND CRITERION . .	128
Hypothesis 2.	128
Hypothesis 3.	135
Hypothesis 4.	138
Hypothesis 5.	143
Hypothesis 6.	145
VIII. SUMMARY, CONCLUSIONS, AND SUGGESTIONS FOR FURTHER RESEARCH.	151
Summary	151
Review of the Study	151
Findings	151
Conclusions and Implications	154
Suggestions for Further Research	157
BIBLIOGRAPHY.	160
APPENDICES	
APPENDIX A. Essay Tests	
APPENDIX B. Objective Tests	
APPENDIX C. Raw Scores on All Tests	

LIST OF TABLES

TABLE		PAGE
I.	Pilot Study Data on Written Expressional Vocabulary	
	Test and on Sentence Sensitivity Test	89
II.	Experience of Raters	104
III.	Ranges, Means, and Standard Deviations for 151 Students	
	on the Criterion Variable and Five Predictor	
	Variables	109
IV.	Means and Standard Deviations of Scores of the Four	
	Readings on each of Three Essay Topics	115
V.	Correlations between the Single Readings of each of Four	
	Readers on each of Three Essay Topics	116
VI.	Correlations between Readings of each of Four Readers	
	on each of Three Pairs of Essays	119
VII.	Correlations between Readings of each of Four Raters	
	on Three Essay Topics	121
VIII.	Table Showing Increases in Correlations with the	
	Addition of Essays and Raters	122
IX.	Summary of Analysis of Variance of Essay Reading.	124
X.	Correlations of Objective Tests with the Sum of Four	
	Scores on each Essay and with the Total Essay Score . .	129
XI.	Multiple Correlations for Combinations of Predictors	
	with Total Essay Score as Criterion	131

TABLE	PAGE
XII. Intercorrelations among Predictor Variables and the Criterion Variable	134
XIII. Contribution of Predictor Variables with Total Essay Score as Criterion	136
XIV. Correlations of Six Predictors with Total Score on a Two-Essay Criterion (Essay One: STEP 2A; and Essay Two: STEP 2D).	140
XV. Multiple Correlations for Selected Combinations of Predictors with a Two-Essay Criterion (Essay One: STEP 2A; and Essay Two: STEP 2D)	141
XVI. Contribution of Predictor Variables with Total Score on Two Essays as Criterion and with Paragraph Used as a Predictor	144
XVII. Contribution of Predictor Variables with Total Essay Score as Criterion	147
XVIII. Contribution of Predictor Variables with Total Score on Two Essays as Criterion	149

LIST OF FIGURES

FIGURES	PAGE
I. Comparison of Means Showing the Rater by Topic	
Interaction	127

CHAPTER I

INTRODUCTION

I. THE DEBATE: OBJECTIVE VS. ESSAY TESTS OF WRITING

In recent years there has been considerable debate among educators, teachers, and psychometricians on the relative merits of objective testing and of actual writing as effective measures of students' competence in written composition. Generally, the supporters of objective tests quote from an extensive body of research and offer the following reasons for their dissatisfaction with theme grading as a measure of students' writing ability.

1. Four variables (as described in Chapter IV) operate to make the reliable assessment of writing ability difficult.
 - (a) A writer variable, which is defined as the inability of a given pupil to perform at a consistent level of overall quality from one writing assignment to the next (Anderson, 1960; Braddock, 1963; Diederich, 1946; Kincaid, 1953; Marton, 1961; and Traxler and Anderson, 1935).
 - (b) An assignment variable which suggests that such factors as the topic, the mode of discourse, the time afforded for writing, and the examination situation will produce variations in the quality of writing (Diederich, 1946; Frogner, 1933; Godshalk, Swineford, and Coffman, 1966).
 - (c) An intra-rater variable, or the tendency of a rater to vary in his own standards of evaluation (Braddock, 1963).
 - (d) An inter-rater variable, or the tendency of several raters to vary from each other in their evaluations. Braddock (1963) quotes several studies to document this variable and points out that agreement between raters may vary from .31 to .96.

2. The controls that research has shown to be necessary, if the evaluation of essays is to be reliable, often become excessively cumbersome and too involved to be of practical use for the average composition teacher, particularly if the number of students being tested is at all large.

Because of this lack of reliability of essay tests as measures of students' writing, educators have been using objective tests of writing for more than thirty years. The proponents of these tests also present results of experimental studies to support their contention that objective measures are reliable, and valid, and economical to score. The following points illustrate their beliefs:

1. Palmer states that many years of College Board English testing have amply proved that "whatever their faults, objective English tests do constitute a reliable and valid method of ascertaining student compositional ability" (1961, p. 34).
2. Godshalk, Swineford, and Coffman conclude that "when objective questions specifically designed to measure writing skills are evaluated against a reliable criterion of writing skills, they prove to be highly valid (1966, p. 40).
3. The defenders of objective tests point to the helpfulness of objective measures as predictors of success in future

composition classes (Braddock, 1963, p. 41) and that objective tests correlate higher with such criteria as course grades in English and teachers' ratings of student writing ability than do essay scores (Palmer, 1961, p. 36).

Supporters of essay tests, however, have difficulty in accepting objective tests as valid measures of writing and generally point to the following as shortcomings of objective tests:

1. Objective tests "do not require the student to select his own words and to compose - to formulate and organize his own ideas into paragraphs and sentences" (Braddock, 1963, pp. 40-41).
2. Objective tests make little or no attempt to measure the "larger elements" of composition, even indirectly (Braddock, 1963, p. 42). Most objective tests emphasize mechanical skills, such as usage and spelling, at the expense of style and quality.
3. Students may "respond correctly to items which they do not use correctly in their own expression" (Greene, 1950, p. 390).

Finally, the supporters of essay tests state that essay tests, by virtue of involving actual writing, are valid by definition and have always had the advantage of apparent "face" validity, in that they require the student to summon and organize his relevant knowledge (Eley, 1955, p. 11). Thus, as Huddleston states, the essay test has been thought of as a "natural" task, allowing a direct approach to

important goals (1954, p. 165). But as far back as the 1880's it was realized that this otherwise ideal testing device was beset with problems of unreliability.

II. THE GENERAL PROBLEM

In view of these conflicting beliefs, what then is the English composition teacher to do about objective tests and actual writing as measures of writing ability? Clearly, the need exists for more carefully designed investigations to find satisfactory and consistent answers to the following questions:

1. Can the unreliability of essay tests be minimized by getting the students to write on several topics and by having these different essays graded by several raters using a holistic marking method? Can these essays scores be used as a reliable criterion against which to validate objective tests of writing?
2. Can one obtain, by selecting from the objective tests available and by constructing tests that may measure writing more directly than has been done, a battery of objective tests that would correlate significantly with a reliable criterion of writing performance? What would be the relative validities of these various objective tests? What combinations of objective tests would have highest validity coefficients?

III. PURPOSES OF THE STUDY

One purpose of this study involved making an attempt to measure writing ability more directly than has been done and to measure more germane elements of writing by constructing and validating two objective tests of writing skills. A Sentence Sensitivity test, measuring several aspects of students' ability to identify and compose well-formed and mature sentences, and a Written Expressional Vocabulary test measuring students' variety and precision in expression were constructed. This decision to test students' knowledge of effective word and sentence choice was made on the basis of the investigator's concept of what students do when they write effectively--a concept derived from the literature, theory, and research relevant to written composition. This explication of writing ability and the writing process served as a comprehensive plan against which the skills selected for testing may be viewed.

A further purpose of this study was to select, in the light of the investigator's formulation of the concept of writing ability and the writing process, three standardized objective tests measuring other germane aspects of writing and to include these with the investigator's two tests to determine the relative validities of the five tests. The three standardized tests selected (the Sequential Tests of Educational Progress: Writing; the Cooperative English Test: English Expression; and the Cooperative English Test: Reading Comprehension) have been evaluated more or less effectively, but on separate

populations. In this study they were included with the two tests constructed by the investigator to form a large-scale investigation in which variables were pooled to determine: (1) the validities of the five objective measures of writing ability as predictors of essay measures of writing performance; (2) the relative validities of selected combinations of the objective tests; and (3) the total variance accounted for by the combination of all predictors and the unique contribution of each predictor to this total variance, with specific attention given to the unique contributions of the Sentence Sensitivity test and the Written Expressional Vocabulary test.

The criterion of writing performance against which the five objective tests were validated consisted of three samples of student writing (two forty-minute essays and one twenty-minute paragraph), with each writing sample graded by each of four raters. The procedure for establishing this criterion is described in Chapter IV.

It was also the writer's purpose to determine the increase in the predictive validity of the selected combinations of objective tests when the twenty-minute paragraph was taken from the criterion measure and included among the predictors. This, however, first necessitated determining the relative validities of the selected combinations of objective tests as predictors of an essay measure of writing performance consisting of the total score on Essays One and Two, each marked by four raters. The total variance accounted for by

the combination of all predictors (including the paragraph) and the unique contribution of each predictor to this total variance was determined. Finally, knowledge of the pupils' intelligence test scores and the sex of the pupil were included with the above predictors and the unique contribution of each of the eight predictors to the total variance was determined.

IV. HYPOTHESES

Relative to the above purposes the following hypotheses were formulated:

1. The total essay score (three essays marked by four judges, or the sum of twelve scores for each student) will have high reading reliability and high total score reliability and can serve as a sound criterion for evaluating the validity of the five objective tests.
2. There will be a significant positive correlation between the total measure of writing performance and the following predictor variables (taken individually and in combinations of two and three):

Written Expressional Vocabulary
Sentence Sensitivity
English Expression
Effectiveness
Mechanics
STEP: Writing
Reading Comprehension
Vocabulary
Comprehension

3. (a) The combination of all five objective tests will account for the highest proportion of total essay score variance, relative to other examined combinations.
- (b) The unique contribution of students' scores on the Sentence Sensitivity Test to this total essay score variance will be significant.
- (c) The unique contribution of students' scores on the Written Expressional Vocabulary Test to this total essay score variance will be significant.
4. (a) There will be a significant positive correlation between scores on a two-essay criterion and the following predictor variables (taken individually and in combinations of two and three):
 - Written Expression Vocabulary
 - Sentence Sensitivity
 - English Expression
 - Effectiveness
 - Mechanics
 - STEP: Writing
 - Reading Comprehension
 - Vocabulary
 - Comprehension
 - A Twenty-minute Paragraph
- (b) The addition of a twenty-minute paragraph as a predictor will increase the multiple correlation between scores on the two-essay criterion and each of the two- and three-objective test combinations.
5. (a) The combination of all five objective tests plus the paragraph will account for the highest proportion of

variance on the two-essay criterion, relative to other combinations.

- (b) The unique contribution of students' scores on the paragraph to this two-essay score variance will be significant.
6. (a) The unique contribution of students' scores on the Sentence Sensitivity Test to the total essay score variance (as in hypothesis 3 b) will remain significant when sex and intelligence are included as predictor variables.
- (b) The unique contribution of students' scores on the Written Expressional Vocabulary Test to the total essay score variance (as in hypothesis 3 c) will remain significant when sex and intelligence are included as predictor variables.
 - (c) The unique contribution of students' scores on the paragraph to the two-essay score variance (as in hypothesis 5 b) will remain significant when sex and intelligence are included as predictor variables.

V. DEFINITIONS

The terms concept of writing ability and the writing process occur frequently in this study. On an operational basis these terms refer to a statement by the investigator designed to approximate as closely as possible the component skills needed by students while writing compositions.

Composition ability (or more specifically composition performance), as the criterion measure, is the sum of four independent judgements on each of three work samples of students' writing. Composition ability is thus defined in terms of the product created by writers with a wide range of skill and the results judged by four qualified English teachers. The sum of twelve marks for each student constitutes a measure of writing performance. Other terms of more limited reference used in connection with the criterion measure, such as reading reliability and reliability of total essay score, are defined at the time of their introduction.

The various predictor variables are defined simply as the students' scores on the appropriate tests: (1) English Expression - including Effectiveness, and Mechanics, (2) Reading Comprehension - including Vocabulary, and Comprehension, (3) STEP Writing, (4) Written Expressional Vocabulary, (5) Sentence Sensitivity, and (6) Intelligence. For example, Written Expressional Vocabulary is the student's total score on the Written Expressional Vocabulary Test constructed by the investigator.

VI. LIMITATIONS

In research in written composition there are so many variables to control that it is an unusual study which does not leave several possibly important variables uncontrolled or undescribed. Though one is free to suspect that students' oral language ability,

their level of motivation, their personality, the teacher variable, and other variables are probably related to students' writing ability, and that these variables would account for some of the variance of scores on the criterion measure, no attempt was made to incorporate these variables into the experimental design. Also, if students were tested on the variables of usage, logic, paragraph organization, outlining, and interlinear exercises and if their scores were inserted into the unrestricted (general) model the amount of variance accounted for by the set of predictors used in this investigation would probably differ, and thus results different from those reported would probably be attained. Additional predictors might well have changed the obtained multiple correlations, or at least have changed the relative weights assigned to the predictors.

A second limitation of this investigation is probably inherent in the criterion, which is not in any sense claimed to be an ultimate criterion measure. As in all essays written under test conditions, the three essays used in this investigation place a premium on fluency and ability to write correctly and with some style in a first draft. In actual life, the writer is seldom under such sharp limitations; he can write and rewrite, use a dictionary or thesaurus, spend long hours outlining, and eventually discover an inherent structure in a complex topic. However, it is assumed in this study that the sample essays collected provide as valid a measure of writing under test conditions as any other that can be obtained in a

similar period of time.

A third limitation is probably associated with the unusual problems of validity and reliability of the two tests constructed by the investigator. The fact that these tests are not standardized imposes a limitation on the study.

Finally, the investigation is limited by the fact that the population involved was entirely from grade ten in only one senior high school in Edmonton, Alberta. All generalizations should be guided by the description of the sample used, as described in Chapter V.

VII. OVERVIEW OF THE INVESTIGATION

This validity study of objective measures of writing ability has been conducted in several phases. These phases are recorded in appropriate chapters.

Chapter II presents the formulation and explication of the investigator's concept of writing ability and the nature of the writing process.

Chapter III reviews the historical development and present status of objective testing of writing ability showing the grounds on which objective tests have been criticized and the challenge facing today's proponents of objective tests. The selection and description of the standardized tests to be used in this investigation and the construction and justification of the two objective tests of Sentence

Sensitivity and Written Expressional Vocabulary are also discussed in this chapter.

Chapter IV reviews the variables associated with essay testing and outlines the procedures followed in establishing a valid and reliable criterion.

Chapter V sets out the research design for the investigation, describes the sample of students, and outlines the statistical analysis used.

Chapter VI presents the analysis of the criterion.

Chapter VII presents the analysis of relations between the objective tests and the criterion.

Chapter VIII, the concluding chapter, contains a summary of the main findings together with certain implications for composition teachers and suggestions for further investigation.

CHAPTER II

CONCEPT OF WRITING ABILITY AND THE WRITING PROCESS

I. INTRODUCTION

It is not enough that objective tests be reliable and correlate highly with a reliable criterion; they ought to be supported by an adequate explication of the concept of writing ability and of the nature of the writing process. This point is well-stated by Steinmann who, in his review of The Measurement of Writing Ability by Godshalk, Swineford and Coffman, states:

Measurement of writing ability presupposes not only knowledge of measurement but knowledge of writing ability. This test may well have a good deal of the former to support it, but neither the test itself nor the criterion against which it is validated seems to have much of the latter (p. 79)....

My central objection to this research is, I think, clear. My objection is that...it lacks conceptual validity: far from contributing to knowledge of writing ability, it is not even supported by any such knowledge. And I say this...because it fails to reflect any serious theoretical concern with this concept...and any research in the relevant disciplines (1967, p. 83).

The task, however, of stating or identifying what the writing process and the concept of writing ability might involve is indeed difficult. Hunting maintains that "our problem is not that the components of 'writing' cannot be counted and measured, because theoretically they can be; our problem is that nobody has yet succeeded in the formidable and terribly difficult task of counting them" (1967, p. 39).

Difficult though it may be to define writing in terms which can be taught and tested, this chapter will attempt to provide an approximation of the concept of writing ability. Such an approximation was used to provide some degree of conceptual or theoretical validity for the selection and construction of objective tests used. The development of this concept proceeds in two stages. First, a review is made of what educators and writers themselves say of the writing process so as to provide support for the formulation by the investigator of a statement, in behavioral terms, of what the student does when he writes effectively. A second stage is to provide a summary of information from disciplines related to written composition to explicate more fully this concept of writing ability and to develop the framework against which to view the five objective tests.

II. STATEMENTS OF LEADING EDUCATORS

Janet Emig in "On Teaching Composition: Some Hypotheses as Definitions," asks: "We teach composition: Whatever can we mean?" "What is the nature of the writing process into which we as teachers intervene?" (1967, p. 130). In partial answer to these questions Emig points out that "In literary, rhetorical, and textbook canon there is a strong tradition that all writers engage in a monolithic process, with that process made up of three discrete components - planning, writing, and revising. Although these canons seldom supply tight or full descriptions for these components, teachers and textbook writers usually agree on the following operational definitions:"

Planning is the sum of those activities, mental and written, the writer engages in prior to producing a first draft.

Writing is his effort to formulate - usually observing the grammatical requirements, semantic conventions, and graphic amenities of a language - an effective expressive or expressive-communicative sequence of words.

Revising is that activity, or series of activities, by which the writer adjusts, at a time usually separated from the writing of a draft, part or all of that draft to more closely approximate certain substantive and stylistic aims (1967, pp. 130-131).

Emig questions whether the writing process is really a fixed and full ordering of these three components occurring in a lockstep, non-recursive, left-to-right sequence and suggests that "instead of a single process of writing there may be processes of writing, at least a process that can be changed - shortened, lengthened, transmorgified - by a number of variables. Instead of a process or processes inexorably made up of three 'stages', there may be more or fewer components. Writing may be recursive, a loop rather than a linear affair - one can write, then plan; or one can revise, then write" (1967, p. 131).

That there is an abundance of support for what Emig calls the "monolithic process" is easily documented by a cursory review of the many textbooks in written composition. "Distinguishable acts" of the writing process are identified in the following representative textbooks:

Smith and Liedlich, From Thought to Theme, 1968, pp. 221-222;
 Fowler, Teaching Language, Composition, and Literature,
 1965, pp. 135-147;
 Cain, Common Sense about Writing, 1967, p. 23; and
 Perrin, Writer's Guide and Index to English, 1965, pp. 38-59.

Two examples will be sufficient to reveal this "process". For Perrin the "Stages in Writing a Paper" are:

1. Focusing on a subject - Definition of topic, sensing problems involved and possible sources of information
2. Gathering material - Notes (in mind or on paper) from memory, observation, interviews, reading, speculation
3. Deciding on methods of development - The ways of approaching and exploring the subject
4. Organizing the paper - A synopsis or outline of the paper
5. Writing the first draft - Tentative copy of the paper
6. Revising - Necessary changes in material, corrections and improvements in words, sentences, paragraphs
7. Preparing the manuscript - The completed paper, ready for reading or for printing
8. Seeing the manuscript into print - The printed copy

Similarly, for Fowler the process of writing is viewed "as a series of six sequential steps". However, Fowler further states that:

Organization, accuracy, clarity, and economy are probably the virtues most in demand in writing today. In addition to these, most thoughtful teachers wish to encourage students to write honestly and responsibly using language with care, integrity, and sensitivity (1965, p. 133).

Focusing more specifically on the concept of writing ability, Martin Steinmann, Jr., in "A Conceptual Review," states:

Writing ability is, I take it, a writer's ability to write effectively, to choose words and word patterns so as to achieve his purpose; his ability is measured by the effectiveness of his choices. To be effective, his choices must take account both of many variables (for example, language or dialect, purpose, reader, occasion or context, emotions to be expressed, and facts to be described) and of the rules or principles relevant to those variables. These rules or principles, and consequently his choices, are (if his language is English) of three broad types:

- Type (1) - grammatical, semantic, and mechanical rules (choices between non-English and English expressions);
- Type (2) - rhetorical rules (choices between synonymous English expressions);
- Type (3) - other rules or principles, logical and psychological, for example (choices between non-synonymous English expressions). (1967, pp. 79-80).

Whereas Emig and Steinmann both stress the importance of the writing process in general, other writers (Rohman, 1965, Guth, 1964, and Wilson and Lacampagne 1967) emphasize "pre-writing" which they deem as an essential first stage in the writing process. Essentially, Rohman defines pre-writing as the stage of discovery in the writing process when the person assimilates his "subject" to himself. It is the stage of the writing process characterized by "thinking", that is, by an "activity of the mind which brings forth and develops ideas, plans, designs...essentially the imposition of pattern upon experience" (Rohman, 1965, p. 106).

Because this stage of the writing process, which the writers call "pre-writing", is within the mind and consequently hidden, it resembles what John Ciardi calls a "groping". In Ciardi's terms what the writer is groping for is that "combination which 'clicks' for him;

an arrangement that will fit his subject to him and him to his subject" (Rohman, 1965, p. 107). Pre-writing, then, is that stage which concerns itself with "discovery" - discovery "not of something at all, but of a pattern of somethings" (Rohman, 1965, p. 107). As Rohman further states:

It is more useful to think of writing not as made up of words but of combinations of words. The meaning of writing is the meaning of the combination, the pattern that the meanings of the many words made when fused by a writer's consciousness in the moment of "discovery" ... "good writing" is that discovered combination of words which allows a person the integrity to dominate his subject with a pattern both fresh and original (Rohman, 1965, p. 107).

For Hans Guth the process of composition is also "first of all a process of exploration". Guth maintains that:

There is an essential preliminary stage of investigation and woolgathering and false starts. There is a gradual collecting of notes - mental and written, concrete and abstract, peripheral and to the point. This stage is followed by an important intermediate one: the sorting out and ordering of final impressions, the seeking out of missing information, the reviewing of evidence in order to test and confirm tentative conclusions. The final result is the statement and support not merely of an honest opinion but of a considered opinion. It is the formulation of judgments more balanced and responsible than we are likely to formulate in the heat of discussion (1964, pp. 165-166).

With writing considered in this manner it is evident that organization (and planning) is not something imposed from without. Instead, organization becomes a matter of doing justice to the subject. As Guth states: "the writer explores his subject in order to discover its inherent structure. He examines his assumptions in order to determine the necessary steps in his argument" (1964, p. 166).

Although theory on such a process as "prewriting" or "precomposition" is not yet complete one can easily agree with Wilson and Lacampagne who, in their summary, "Developments in Composition," state that "virtually all of the promising programs in composition stress the more positive procedure, the precomposition process (1967, p. 63).

Closely related to the "pre-writing" stage of the writing process, if not identical to it, are the many comments (Loban, 1961, Kitzhaber, 1965, and others) stressing the importance of "clear thinking" as a requirement for effective writing.

To write clearly students must think clearly, to write competently, they must think competently, to write with power or imagination, they must think with power or imagination. Think...write...write...think... these processes cannot be disjoined. When a student has learned to write better he has learned to think better. This is a law. There is no way around, only through (Loban, Ryan, Squire, 1961, p. 485).

The precomposition process, however, goes well beyond the selection of a subject, clear thinking, and the development of ideas, for students at the pre-writing stage must find not only a general purpose for communicating but also a special purpose for writing about a particular subject. Kitzhaber makes the point by stating:

By good writing I mean also writing that shows consciousness of a purpose, and of a reader who is being written for. Writing that neglects purpose, and takes little or no account of the reader is aimless, unfocussed, and almost invariably dull. Finally, by good writing I mean writing that is responsible, writing that shows an awareness of the right of the reader, that treats him with sincerity and respect (1965, p. 112).

III. WRITERS ON WRITING

Writers, too, have commented on writing ability and the writing process. Malcolm Cowley in his "Introduction" to Writers at Work: The Paris Review Interviews states:

There would seem to be four stages in the composition of a story. First comes the germ of the story, then a period of more or less conscious meditation, then the first draft, and finally the revision, which may be simply "pencil work", as John O'Hara calls it - that is, minor changes in wording - or may lead to writing several drafts and what amounts to a new work (1967, p. 7).

Similarly, Brewster Ghiselin, in his "Introduction" to The Creative Process, states that:

interest in the creative process is not exactly a new development...both Plato and Aristotle had something to say on the creative process, and from time to time during the next two thousand years other writers touched upon it...Blake, Wordsworth, Coleridge, Shelley, and Keats all had their say... Interest in the subject is still growing (1961, p. 11).

But what does such a collection of writings, as edited by Ghiselin, suggest that can be useful in the formulation of a general statement on the writing process? The following quotations are explanatory in this connection:

Not even the most vigorously creative minds always find their way quickly to efficiency (p. 11).

Creation begins typically with a vague, even a confused excitement, some sort of yearning, hunch, or other preverbal intimation of approaching or potential resolution. Stephen Spender's expression is exact: "a dim cloud of an idea which I feel must be condensed into a shower of words." Alfred North Whitehead speaks of "the state of imaginative muddled suspense which precedes successful inductive generalization," and there is much other testimony to the same effect (p. 14).

Production by a process of purely conscious calculation seems never to occur...it does not fit the facts reported almost universally and in every field of creative work (p. 15).

...the process is an organic development and it helps to dispel the notion that creation is simply an act of canny calculation governed by wish, will, and expediency (p. 21)

Among the conditions to which every inventor must submit is the necessity for patience (p. 26).

Plan must come as a part of the organic development of a project, either before the details are determined, which is more convenient, or in the midst of their production, which is sometimes confusing (p. 27).

Two important stages in (the creative process) are predominantly conscious and critical, and in these the will properly functions. It is of use in that preliminary labor, or sometimes less burdensome preparation, without which there can be no significant activity, and in the work of verification, correction, or revision that ordinarily follows the more radical inventive activity and completes or refines its product (p. 28).

From the above quotations and from the specific selections in The Creative Process, one observes a lively sense of the divergencies of individual approach and procedure. Writing is a highly individual pursuit and thus it is somewhat natural to expect

that the reactions of professional writers themselves to questions on the writing process and the concept of writing ability would reflect a variety of educational philosophies. It is just such a variety that Frank Lowe records in his essay "Writers on Learning to Write," namely, "Miss Buck stresses freedom to create; Mr. Fadiman strongly favors the 'traditional' approach and curriculum; Dr. Spock emphasizes the importance of environment, especially the home" (1964, p. 494). More specifically, of the eleven questions asked by Mr. Lowe of the fourteen professional writers on how they learned to write, questions eight and nine elicit answers very germane to this study. Lowe records the following:

Question 8: What has been the greatest influence on your writing ability?

Buck: Reading great books.

Chase: A natural curiosity to understand a given situation - war, depression, labor, management, automation, semantics - and then to explain it as lucidly as possible.

Cousins: Meeting many different kinds of people and conceiving a desire to tell others about them.

Golden: Reading other writers.

Hersey: Not what, but who, King James' committee of translators, Tolstoi, Cervantes, Melville, Tsao Hsueh-chin, Stendhal, Conrad; and teachers...

Michener: Wide reading of people who can write better than I can.

Mayer: Reading. Especially my own writing with distaste.

Spock: (Aside from the general literacy level of family), my mother insisted, during any absence from home, that we write vivid, detailed descriptions of all happenings, people.

Question 9: What effect did reading have on your interest in writing?

Buck: Profound effect.

Buckley: Considerable.

Chase: Immense!...

Cousins: Reading had an immense effect on my writing: it expanded my horizons, put me in touch with minds I could emulate and interact with despite barriers of time and place.

- Fadiman: Obviously a considerable one - writers read, don't they?
- Gann: It is one of the most important things any writer can do.
- Golden: Everything depends on my reading.
- Mayer: Enormous...
- Shaw: Reading is the beginning of any writer's career, long before he has any notion he wants to write.
- Spock: I read a lot and assume it sets tastes and created identification with writers (1964, pp. 491-492).

One conclusion that can perhaps be drawn from these confessions of the power of good books is that through reading writers not only develop a reservoir of ideas from which to draw, but that wide reading enables writers to confirm and clarify their experiences by comparing them with those of others and at the same time enables them to become acquainted with precise, forceful, and stimulating expression. As Guth states, "the student must see the strategies and devices of good prose at work in a wide variety of reading, assigned and unassigned" (1964, p. 173).

IV. GENERAL STATEMENT OF CONCEPT OF WRITING ABILITY AND WRITING PROCESS

In general, a synthesis of the above statements of educators and writers suggests that the following are perhaps distinguishable acts or components of writing.

- I. The topic is assigned to the student
- II. The student goes through a process of discovering the subject context, that is of formulating a theme based upon the topic or idea

A. There is a preliminary stage when the student:

1. Considers the topic
2. Considers his purpose in writing (the reason for his communication)
3. Considers his reader or audience
4. Considers the uniqueness of the topic
5. Considers the sources
6. Investigates, "gropes", makes false starts, and gradually collects notes (mental and written, concrete and abstract, peripheral and to the point)
7. Reads, thinks, discusses
8. Delimits subject in accordance with purpose
9. Formulates a central theme

B. The student selects the material which is necessary for the development of his theme:

1. He draws upon his own resources, ideas, experiences, thoughts, knowledge, impressions, etc.
2. He consults other sources
3. He evaluates his material in the light of his central theme and purpose
4. He discriminates between major and minor points
5. He sorts out and orders his impressions, he makes a tentative outline
6. He considers introduction to the central theme
7. He views evidence to test and confirm tentative conclusions
8. He places points in an order appropriate to his purpose and material
9. He considers his concluding points

C. The student proceeds with the actual writing; that is he makes his statements and support of an honest and considered opinion, he makes his formulation of judgements, balanced and responsible.

III. In the actual writing process, the effectiveness of the student's finished product will depend to a considerable degree upon his ability to make the best choices possible from the many alternatives available. This selective process might be summarized as follows:

A. From the vast resources of the English Language, he must choose:

1. words,
2. word patterns,
3. sentence patterns and variations, and
4. paragraph structures.

- B. His choices of words, patterns, and structures must be related to:
 - 1. his purpose in writing,
 - 2. the reader,
 - 3. the facts to be described,
 - 4. the emotions to be expressed, and
 - 5. appropriate usage.
- C. The effectiveness of his choices will be affected by:
 - 1. stylistic features and psychological rules,
 - 2. rhetorical considerations,
 - 3. logical rules,
 - 4. semantic relationships,
 - 5. awareness of form,
 - 6. grammatical rules,
 - 7. conventional requirements, and
 - 8. aspects of personal integrity.

IV. The student must know and use the techniques of proof-reading and revising.

One notes that in Sections One and Two the student is concerned with discovering what Rohman (1965) would call the "subject context". It is a period when the student "assimilates" his "subject" to himself. It is an active period characterized by "thinking", that is with "thinking" considered as an activity of the mind which brings forth and develops ideas, plans, and designs. In Section III, or the process of struggling and of imposing pattern upon experience, the student has to discover a "personal context"; that special combination of words that makes an essay his, not yours or mine. The student has to choose words and combinations of words to convey his meaning. The meaning of writing is the meaning of the combinations, the pattern that the meaning of many words make when fused by a writer's consciousness in the moment of "discovery" (Rohman, 1965, p. 107).

In the writing process the writer gropes for words and sentences that will reveal his experience. But, as suggested in the general statement above, these choices must be made on the basis of several main variables which in themselves are supported by disciplines related to written composition. For example, from linguistics, psychology, psycholinguistics, logic, semantics, and rhetoric one can identify specific information and techniques that are useful in the teaching and testing of written composition and that can contribute to a further explication of the variables listed above in Category C of Section III. In some elementary form, specific concepts - for example, concepts from semantics - can be made available to a person approaching a writing problem without turning the teaching of writing into a course on Korzybski and perhaps without involving students with much terminology of the particular field.

IV. INFORMATION FROM RELATED DISCIPLINES

Semantics and Written Composition

Although a detailed study of semantics may not necessarily belong in a high school course in English composition, a number of writers emphasize that many semantic concepts can be used in helping students write more effectively, particularly the semantic activities of clarifying abstractions, understanding metaphor, and recognizing connotative language. Semantics can challenge a writer to rethink

his position instead of immediately entrenching himself behind fixed ideas. Lecky says of semantics that:

It can help him (the student) to recognize disputes over words. It can show him that language and thought are so closely interconnected that phrasing is more than a matter of style, to be attended to after the inventing and disposing have been completed; and that metaphor is more than a pleasing departure from familiar language (1967, p. 10).

Semantic theory can be constructively helpful to a writer faced with the task of defining terms. The student of semantics will know that he is defining words (according to usage and in the light of referents), not defining referents. The student will thus be alert to the pull of arbitrary definition and the force of context.

Attention to context, whether the student is reading or writing, is imperative:

A sentence like "People get hurt" has descriptive form, but under certain circumstances its message predicts, threatens, and prescribes. More often a writer must depend on the interplay of words to control meaning. The immediate verbal context, he sees, is of the greatest importance. Delicate connotative and denotative variations keep words from being strict synonyms or antonyms (Lecky, 1967, p. 11).

Surely, an awareness of the above semantic activities, can help the composition student not only to increase the size of his vocabulary, but it can lead to an important improvement in the quality of his discernment, a laudable aim in any composition course.

More specifically, the study of semantics would make the following concepts available to a student approaching a writing situation:

1. clarifying abstractions
2. understanding metaphor
3. understanding connotative language
4. understanding denotative use of words
5. classifying
6. using deductive language
7. being expressive - having resourcefulness and flexibility of vocabulary
8. using apt, precise and vigorous expressions
9. "feeling" for words
10. being alert to distinctions, implications and overtones

Logic and Written Composition

At its best, logic is a subtle science, and efforts to make it part of formal instruction in English often leave both teacher and student bewildered, perhaps frustrated. English teachers are not trained logicians. However, whenever English teachers insist on "clear thinking" in writing they do teach some of the elements of logic informally. But, as Kitzhaber points out:

It should be helpful to student and teacher alike, in a course dealing with expository writing, to identify a few of these principles of clear thinking so they become a conscious part of the student's equipment for analyzing the writing and speaking of others and for guiding his own practice (1963, p. 140).

Logic, as a science, focuses on principles of definition, the nature of evidence, and the main outlines of inductive argument - principles well worth making known to the student to increase his effectiveness as a writer of expository prose. As the Commission on Freedom and Discipline in English states:

It is reasonable, then, for the English teacher to commit himself only to those logical problems most often encountered in discursive writing: the broad differences between inductive and deductive argument, the sense in which all rhetorical arguments are contingent, the usefulness but final invalidity of arguments by analogy, and - most common of all - the slipperiness of enthymeme or argument in which an essential part is omitted because it is assumed rather than stated. Again, professional expository essays and students' papers will suffice as texts if the teacher knows clearly the few matters of logic he wants to teach, for most human discourse consists of propositions so arranged as to be subject to rules of logic (1965, p. 105).

Though logic, outside philosophy and mathematics, receives but a cursory treatment, it does present some elementary ideas to illumine the teaching and practice of certain kinds of writing. Especially helpful is the use of symbolic logic in clarifying the meaning of certain commonly used words as "if", "only", "and", "or", "necessary", and "therefore" (Pitt and Leavenworth, 1968, p. 2). Such words exercise powerful control over the meaning of the sentences in which they occur, and their proper use is crucial to the validity of arguments. Logical form influences the meaning of many sentences and unless students understand how such words determine the logical form of both sentences and arguments they may have difficulty in understanding what such sentences mean and whether their arguments are valid.

Elements of logic are undoubtedly necessary for the correct expression of tentative thinking through the use of provisional and conditional statements. In Loban's study of the language development of elementary school children, expression of tentativeness proved to be a key function of language which distinguished effective and ineffective users of language.

Those subjects who proved to have the greatest power over language - by every measure that could be applied, not just by the combined Teachers' Rating Scale and Vocabulary Test - were the subjects who most frequently used language to express tentativeness. Supposition, hypotheses, and conditional statements occur much less frequently in the language of subjects lacking skill in language (Loban, 1963, p. 53).

In summary, the study of logic would help the composition student to:

1. distinguish particulars from generals;
2. understand and recognize the basic kinds of relationships possible among particulars, and between particulars and generals, especially the relationships of:
 - a. time,
 - b. space,
 - c. similarity and dissimilarity; and
3. to have some idea of the number and kind ("quantification" and "typicality") of particulars needed to develop selected topics.

Psychology, Psycholinguistics and Written Composition

J. B. Carroll points out that there is a critical need in the psychology of language for a comprehensive theory of language behavior. "We need to construct and complete a theory of how people learn to produce and understand language, both spoken and written" (Carroll, 1962, p. 23). B. F. Skinner proposed such a theory in his book entitled Verbal Behavior (1957). His theory was based on a simple and powerful idea, namely that there is a kind of behavior, which is called operant behavior, represented by any response whose probability of emission is controlled by the kinds of consequences it has. That is, the likelihood that a particular response will recur

is contingent upon whether that response brings consequences which are in some sense satisfying or rewarding to the individual. However, Skinner's theory of verbal behavior raises some problems, as is evidenced in the critiques by Chomsky (1959) and Carroll (1962).

Although Skinner attempted to be comprehensive and to explain a large range of phenomena, one can nevertheless see some application of the theory to the field of English composition if one focuses attention on the manner in which verbal behavior is reinforced by various kinds of rewards. Given Skinner's suggestion that behavior is created and altered in an individual by the consequences it produces, a major task of the composition teacher should be to see that his students' academic behavior produces consequences, or reinforcement. Thus, if writing skills are to be reinforced by teacher marks and comments, only favourable comments will do the job, and the more specific the comments are the clearer will be the writing. One could also conclude from the theory of reinforcement that to the extent that teacher's comments are punishing they can be expected to produce some of the complex effects of punishment observed inside and outside the psychological laboratory.

In "The Behavioral Repertoire of Writing" Douglas Porter of the Harvard Graduate School of Education indicates that there are three main aspects of writing:

One is the substantive content of the writing production - the physics, personal history, or imagery, and so forth. The other factors are the language repertoire in the sense of "contentless" techniques...and finally, the discriminations that set the appropriate occasions for the use of one technique as opposed to another (1962, p. 17).

Porter suggests that rather than require the student to struggle with ideas, style discriminations, and the discovery of techniques concurrently, the teacher should deal with these three components of writing behavior individually. He believes that behaviors to be taught effectively must be clearly identifiable and clearly specified so teachers can reinforce them when appropriate. However, it has been suggested by Carroll that this line of thinking probably has limited value when we face the many problems we have in teaching English composition; "It would imply that the teacher would have to arrange for the specific rewarding of the thousands of particular responses which we want to teach" (1962, p. 24).

Even granting that reinforcement (reward), or its apparent opposite, punishment, can have the effect of maintaining or suppressing particular kinds of behaviors, this principle does not necessarily explain how behavior is produced (Carroll, 1962, p. 24). Operationally and experimentally, Skinner has to wait for behavior to occur, whereupon he arranges for it to be reinforced or rewarded in some way. Perhaps, as psychology and linguistics work more closely together, and as linguistic science continues to expand, an increased understanding of composition (seen as a complex behavioral repertoire) will be forthcoming.

One such attempt in viewing composition in relation to the psychology of language is that of Carroll in "Psycholinguistics and the Teaching of English Composition". To the extent that it is one of the chief tasks of psycholinguistics to describe and explain what happens in verbal behavior, Carroll views the teaching of English composition as consisting of seven components: (1) having something to say; (2) gauging the audience; (3) organizing one's thoughts; (4) choosing the right words; (5) constructing sentences and paragraphs; (6) saying exactly what you mean; and (7) saying it with style (1956, p. 188). Carroll suggests that "communication implies an audience, and that lacking an audience verbal behavior is stilted and unnatural" (1956, p. 189). Also significant are Carroll's comments on the third component of writing (organizing one's thoughts):

Good writing shows a peculiar kind of interconnect-
edness which carries the reader along and allows him
to fall into no traps unawares. As you very well
know, this is accomplished by establishing in almost
every sentence some expectancy of what sort of thing
may come next, - perhaps an illustration, perhaps
a contrasting idea, perhaps an explication of a
novel idea just introduced. There is even the trick
of establishing that "nothing" will come next (1956,
p. 190).

From a study of the psychology of language, then, the written composition student would become conscious of style discrimination, appropriate wording, organization, and largely techniques similar to those to be gained through a study of rhetoric.

Rhetoric and Written Composition

While there is considerable academic discussion concerning the boundaries that constitute "rhetoric" and the value of the "old" rhetoric versus the "new" rhetoric, there is general agreement on the relevance of rhetorical concepts to composition programs. The many different articles and texts on rhetoric seem to agree that in some way or in some sense the province of rhetoric is the art of effective expression. At the core of rhetoric there is still the Aristotelian definition of the ways "to convince or persuade an audience to think in a certain way and act in a certain way" (Wilson and Lacampagne, 1967, p. 60). But as a study of how to compose effectively, rhetoric inevitably includes principles from logic, semantics, and linguistics, including the study of usage.

To place rhetoric in some perspective, Martin Steinmann, Jr. points out that a writer must have knowledge of three things if he is to use language effectively:

First, to speak or write at all, one must know English; that is, know how to choose between English and non-English expressions. To know English is to possess the ability to make one's utterances conform to a set of rules - grammatical (concerned with form) and semantic (concerned with meaning and including the rules of deductive logic) - that decide for a given expression whether it is English.

Second, to speak or write effectively, one must know how to think effectively; that is, know how to choose wisely between things to say, between nonsynonymous expressions.

Third, to speak or write effectively, one must know rhetoric; that is, how to choose wisely between ways of saying the same thing, between synonymous expressions (1967, pp. 19-20).

Rhetoric is concerned with the effective choice of synonymous expressions; but as the word "effective" suggests, it is concerned, not with utterances only, the mere forms, but with some of their relations to other things. These other things, to quote Martin Steinmann, Jr., are among the six variables that every act of speech or writing has:

The speaker or writer, his utterance, his context (occasion or medium), his audience (listener or reader), his purpose (the effect that he intends his utterance to have upon his audience), and the effect of his utterance upon his audience (1966, p.22).

Rhetoric is best characterized by reference to these variables and the importance of choosing effectively in accordance with these variables.

A further understanding of the role of rhetoric in an English composition course is implicit in a rather compact statement by Wayne Booth:

The common ingredient that I find in all of the writing I admire - excluding for now novels, plays and poems - is something that I shall reluctantly call the rhetorical stance, a stance which depends on discovering and maintaining in any writing situation a proper balance among the three elements that are at work in any communicative effort: the available arguments about the subject itself, the interests and peculiarities of the audience, and the voice of the implied character, of the speaker (1967, p. 184).

The elements of "subject," "audience," and "speaker", as identified by Booth, are in effect, those emphasized by Aristotle and are found in many modern textbooks on rhetoric, as for example in The Logic and Rhetoric of Exposition by Martin and Ohmann (1963).

Martin and Ohmann discuss the following as broad categories of rhetoric:

- (a) the writer: character, persona, voice;
- (b) the reader: identity, susceptibility, resistance;
- (c) the subject: invention, arrangement, style (1963, pp. 132-160).

It is however with "style," one of the three subdivisions in the "subject" category of rhetoric, that one finds perhaps the richest current discussion of composition. But "style", in this context, is difficult to distinguish as a discrete area of study. Style is a highly complex phenomenon which can be viewed from many points of view. In this study it will be considered with "Linguistics and Written Composition".

Generally, then, from a study of rhetoric the student of written composition learns of the importance of effective word and sentence choices. Particularly, too, does the student learn:

1. to show consciousness of purpose in his writing;
2. to show consciousness of a reader, to show an awareness of the rights of the reader and to treat him with sincerity and respect; and
3. to be aware of the strategies of rhetoric, that is the requirements for securing the reader's attention and consent, the dynamics of persuasion, and the resources of style.

Linguistics (Grammar) and Written Composition

Within the last fifteen years the English profession has witnessed rapid developments in the study of the structure of the English language. Developments in structural grammar and more

recently in transformational grammar have brought new theory and some new evidence to bear on the "traditional" grammar-composition problem. The hypothesis that the logical structure of a generative grammar is analogous to the psychological structure of the process of sentence production is becoming both speculatively and empirically compelling. To explore this hypothesis one would have in effect to explain the theoretical character of generative grammar and its explanatory and predictive power of the native speaker's capacity to produce an infinite number of the sentences of his language. Although such a task is beyond the scope of this study, the writer refers the reader to the work of Johnson (1963, 1964) and Jenkins (1961) who provide empirical evidence of the psychological reality of generative grammar theory. Katz, (1964) too, expounds on the psychological reality of generative grammar theory. In effect, the transformationalists are concerned with discovering a mental reality underlying the data of the speaker's verbal behavior.

In relating this theory to written composition there is at present a guarded optimism and some support for the assumption that "modern linguistics" can lead to the improvement of writing sentences and paragraphs. This optimism for generative grammar is certainly reflected in the work of Bateman and Zidonis:

With the advent of generative grammar, which is in essence a representation of the psychological process of producing sentences, a logical approach to the study of composition has become available.... The composition teacher has always strived to get his pupils to write better sentences, but he has no procedure whatsoever for explaining to them just what the concept sentence contains. A pupil who has only a vague notion of sentencehood is doubtless at a disadvantage in evaluating the quality of the sentences he has produced or in understanding the constructive criticism of them offered by his teacher. Conscious control of well-formed sentences seems fundamental to the act of writing, but what is not understood can hardly be controlled. Pupils must be taught a system that accounts for well-formed sentences before they can be expected to produce more of such sentences for themselves. It is the function of a complete grammatical system to define the concept of sentencehood... (1966, p. 2).

In general, one finds in linguistic studies and discussions of composition, the idea that "new grammar" has supplied a principle vital to the art of writing. It is a principle that has been formulated partly because modern linguistics has lent itself particularly well to inductive methods of teaching - that is, methods which present the study of language as a process of discovery, methods which stem from the assumption that language has patterns that are within the students' powers to observe and analyze and that these patterns can be taught along with a great variety of structures of modification for these patterns.

Other current linguistic discussions pertaining to written composition are found in the work of Francis Christensen (1967, pp. 190-199; 1967, pp. 200-216). Under the title of "rhetoric" and

in conjunction with the word "generative", Christensen discusses the idea of rhetoric as generative and the possibilities of teaching "the sentence as professionals use it" (1967, p. xii). This use of the term generative, it appears, results from the feeling that as generative grammar seems closest to the psychological process of sentence formation, so a generative rhetoric is closest to the creative process. What emerges from this generative approach is a body of principles - not prescriptions - which the student may use creatively to generate sentences and paragraphs to suit best his own purpose in informing, persuading, or moving his reader. Under the guidance of the teacher the student can discover the principles operating in models of writing and learn the possibilities of order available to him (Wilson and Lacampagne, 1967, p. 64).

VI. FURTHER EXPLICATION OF CONCEPT OF WRITING

At this point it is possible, in view of the brief review of some areas related to written composition, to provide a further explication of the concept of writing ability as outlined on pages 25-26. It is particularly necessary to elaborate on Category C of Section III in such a manner as to facilitate a rough matching of some components of writing with the skills tested in the five objective tests used. However, because of the interrelatedness of the subject areas basic to written composition, and because of the complexity of the many elements of effective writing not accounted for by these subject areas, many parts of the following elaboration are

stated as germane to writing though justification for their inclusion is not presented. The following, then, govern, in large measure, the students' decisions and choices relevant to particular words, word patterns, sentence structures and variations, and paragraph organization.

Stylistic features and psychological rules

This variable may best be described as a combination or integration of the following that results in something special and distinctive for the student:

1. Sentence structure - varied and smooth
2. Diction - vivid and suitable
3. Figurative language - fresh and fitting
4. Tone appropriate to purpose and subject
5. Freshness of approach
6. Revelation of personality and expression of mood
7. Use of local color and dialect
8. Characterization through dialogue
9. Humor or surprise twist (anti-climax, understatement, exaggeration, mock-seriousness, pretentious language for effect)
10. Effective use of devices such as examples, illustrations, analogy, allusion, alliteration, assonance, rhythm
11. Effective touches of realism; sensory impressions aptly used

Rhetorical considerations

1. The student shows consciousness of purpose
2. The student shows consciousness of reader, he shows an awareness of the rights of the reader and treats him with sincerity and respect
3. The student is aware of the strategies of rhetoric, that is, the requirements for securing the reader's attention and consent, the dynamics of persuasion, and the resources of style

Logical rules

1. Students should be able to distinguish particulars from generals

2. Students should be able to understand and recognize the basic kinds of relationships possible among particulars, and between particulars and generals, especially the relationships of:
 - a. time
 - b. space
 - c. cause and effect
 - d. similarity and dissimilarity
3. Students should have some idea of the number and kind ("quantification" and "typicality") of particulars needed to develop selected topics.

Semantic relationships

1. Clarifying abstractions
2. Understanding metaphor
3. Understanding connotative language
4. Understanding denotative use of words
5. Classifying
6. Using deductive language
7. Being expressive - having resourcefulness and flexibility of vocabulary
8. Using apt, precise, and vigorous expressions
9. "Feeling" for words
10. Being alert to distinctions, implications and overtones

Awareness of form

An understanding of this variable may best be indicated by answers to the following questions:

1. Does the writer explore his subject in order to discover an inherent structure?
2. Is the writer aware of how form involves organization in the sentence, in the paragraph, and in the essay as a whole?
3. Does the introduction prepare the reader for what follows?
4. Is it possible to state clearly the central idea of the essay?
5. Is there a clear relationship among main ideas? Are they developed in a logical order?
6. Are transitions from one idea or topic to another clearly made?
7. Are individual paragraphs sufficiently
8. Is the central idea of the paper as a whole sufficiently developed through the use of details and examples? Does the writer make clear where he stands and does he support what he has to say?

9. Does the theme have a definite satisfactory conclusion?

Grammatical rules

1. Are the sentences grammatically correct? (For example, is the writing reasonably free of fragments, run-on sentences, comma splices, faulty parallel structure, mixed constructions, dangling modifiers, and errors of agreement, case, and verb forms?)

Conventional requirements

1. Appearance and general format (spacing of title, margins, etc.)
2. Documentation (footnotes, bibliography)
3. Spelling
4. Punctuation
5. Capitalization
6. Abbreviations
7. Numbers

Aspects of personal integrity

1. Conviction - The student's writing must be something that he is willing to stand up for, something to which in some way he is personally committed. Writing with conviction means:
 - a. recording one's own observations
 - b. pinning down one's own reactions
 - c. interpreting one's own experience
 - d. formulating one's own judgements
 - e. questioning one's own premises
 - f. making one's own mistakes
2. Good writing involves responsibility. A writer has an obligation to his own standards of fairness and of truth.

VI. SUMMARY

This chapter endeavored to explain and to support a concept of writing ability and the nature of the writing process derived by the investigator on the basis of a review of pertinent statements by educators and writers, and on the basis of a brief review of subject

areas related to written composition. The statement is thus a formulation of what students do when they write effectively. Essentially, it was stated that shaping a composition involves choices. The nature of these many choices and their interdependence was explained.

It is on the basis of this explication of writing ability and the writing process that the three objective tests selected for inclusion in this study and the two objective tests constructed by the investigator are evaluated in the next chapter.

CHAPTER III

OBJECTIVE TESTS

I. REVIEW: OBJECTIVE TESTS OF WRITING ABILITY

When objective tests in English first began to appear they were not so much designed for purposes of measuring achievement as for a means of diagnosing difficulties in grammar, punctuation, and usage (Enochs, 1948, p. 27). This use of objective tests as diagnostic instruments was in its beginning closely related to numerous error studies conducted in the years following 1915. Lyman in his Summary (1929) has presented such a thorough review of the error studies made and the objective language tests constructed and used up to 1929, that rather than repeat what has already been done so well the investigator refers the reader to that study for details. However, from Lyman's description and summary, two observations are pertinent: (1) the objective tests which Lyman reviewed measured only one of the objectives of writing, that concerned with grammatical correctness, and (2) they were designed to diagnostic purposes rather than to measure achievement in writing.

The various developments in the construction of objective tests in English and the many validity studies conducted during the years 1929-1948 have, by and large, been discussed by Enoch in

Measuring the Ability to Write (1948, pp. 27-44). After evaluating ten objective tests on the basis of questions on their validity, reliability, and limited range of writing objective, Enochs states the following general conclusion:

Some tests of the short answer type were still subjective in scoring, others were difficult to score, and most of them were found to be very limited as to the objectives they tried to test. Techniques, too, were not much related to the realities of the process of writing, nor was content particularly well adapted to the maturity level of those persons who it was known would be taking the writing tests (1948, p. 44).

Enochs also provides a thorough review of the beginning efforts of the Cooperative Test Division of the American Council on Education. He describes the Cooperative English Test: Usage, Spelling, and Vocabulary of 1932, the 1937-39 efforts at adapting the tests to machine scoring, the early subsequent validity studies and their leading to a new series of tests: Test A: Mechanics of Expression, and Test B: Effectiveness of Expression. Generally, the test-makers of the Cooperative English Tests attempted to analyze the process of writing more thoroughly and to measure more of the phases of writing than have been reported in earlier tests. As the manual accompanying the tests points out:

In preparing the outline (for Test B: Effectiveness of Expression), an analysis was made of various factors contributing to effectiveness in the use of English. Various English courses of study, composition textbooks, and general discussions on the problems met in writing effectively were used as aids in the analysis (Enochs, 1948, p. 41).

The foregoing review indicates that the problem of measuring writing ability in written composition through the use of objective tests has largely been attacked in a piecemeal manner. Tests with varying degrees of promise have been evaluated more or less effectively, but on separate populations. It was Huddleston (1954) who attempted a large-scale investigation by pooling variables to determine their interrelationships. Huddleston compared the validities of the verbal sections of the Scholastic Aptitude Test (SAT) and of an objective editing test with those for essay ratings. Her general conclusion was that:

The verbal test is more closely related to writing ability as defined in this study (correlations of .76 and .77) than is any other variable. The other variables, when combined in a multiple regression equation with the verbal test, fail to add appreciably to the relationship to writing ability demonstrated by the verbal test alone (1954, p. 204).

The objective editing exercise, however, did raise the correlations .03 and Huddleston concluded that a combination of SAT-verbal scores and objective English scores provided the best prediction available at that time.

Other studies in the 1950-60 period provide further evidence of validity for objective and semi-objective tests, while the testing techniques themselves were being somewhat more refined. Godshalk, Swineford, and Coffman provide the following summary of these efforts:

Swineford and Olsen (1953) reported satisfactory reliability and validity for the interlinear exercise.... Coffman and Papachristou (1955) reported preliminary success with a number of questions designed on the basis of an analysis of student essays. Thomas (1956) succeeded in adapting Stalnaker's construction shift question to a machine scorable format. Miller (1953) pointed out the great variety of skills tested by the interlinear exercise. Weiss (1957) demonstrated that several of the types of questions used in the College Board English Composition Test were making distinctive contributions to the prediction of College English grades (1966, p. 3).

Perhaps the most successful study of the reliability and validity of various types of objective questions designed to measure writing ability is the study by Godshalk, Swineford, and Coffman, described in the monograph, The Measurement of Writing Ability. In this study, the authors examined the relationships among ratings of five themes written by each of 646 secondary school juniors and seniors and the students' scores on the following eight tests:

- (1) Usage items require the recognition of faulty usage, including grammar, diction, basic structure, and mechanics.
- (2) Sentence correction items require the selection of a best form for an underlined portion of a sentence.
- (3) The paragraph organization type requires the student to reassemble several "scrambled" sentences into a coherent paragraph and to show the correct order of sentences.
- (4) Prose groups consists of groups of items, each based upon a paragraph with a sentence omitted. For each of four sentences that might supply the omission, the student is required to judge whether it is appropriate, deficient in sense, inadequate in tone or diction, or deficient because of grammar or construction.
- (5) Error recognition items require the student to decide whether a sentence is free from error of the kinds described, or if not, to indicate which of four classes of error it illustrates: faulty diction, verbosity or redundancy, cliches or abused metaphors, or faulty grammar or structure.

- (6) The construction shift item requires the student to decide what additional changes to make in a sentence if a specified element is changed in a certain way.
- (7) and (8) Two interlinear exercises consisting of poorly written material that requires the student to find and correct deficiencies. (1966, pp. 6-7)

The obtained correlation coefficients for the various objective tests with the criterion (the sum of the ratings of each of five readers on each of five themes) were for test (1) .707, for (2) .705, for (3) .458, for (4) .568, for (5) .592, for (6) .645, while the two interlinear exercises had coefficients of .644 and .668.

Validity coefficients of selected combinations of subtest scores taken three at a time were calculated to find the most effective and efficient set of predictors of writing skill, shown by the criterion measure. These coefficients ranged from .706 (variables 3,4, and 8) to .767 (variables 1,6, and 8).

Later, Godshalk et. al. used one of the twenty-minute essays as a predictor, thus reducing the criterion measure to the sum of marks on four essays. Different combinations of, either one of the interlinear exercises or one of the essays, with some of the objective tests plus the verbal score from the PSAT or SAT were formed. It was found that in all cases a combination including an essay resulted in a higher multiple correlation coefficient with the four essay criterion than did the same combination with an interlinear exercise replacing the essay (.733 to .796 as compared to .712 to .757).

In a most succinct statement, Edward Noyes has summarized the main findings of the Godshalk study as follows:

It is enough to say that, checked against a criterion far more reliable than the usual criteria of teachers' ratings or school and college grades, all but one of the item types currently used in the English Composition Test proved to be excellent predictors; that a very high correlation was achieved when, for a typical one-hour test, two objective item types were combined with an interlinear exercise; and that a 20-minute essay -- read, not analytically, but impressionistically and independently by three readers -- contributed somewhat more than even the interlinear exercise to the validity of the total score. The combination of objective items (which measure accurately some skills involved in writing) with an essay (which measures directly, if somewhat less accurately the writing itself) proved to be more valid than either type of item alone.

In addition to the College Board English Composition Test other standardized objective tests of writing skills also claim to measure varied aspects of writing ability and the writing process. In particular, three such tests deserve mention:

1. Sequential Tests of Educational Process (STEP):
Writing by Educational Testing Service
 - A. Organization: Reasonable ordering of ideas, events, facts, etc. (The ability to organize materials: in the whole passage, in the paragraph, and in the sentence.)
 - B. Conventions: Attention to syntax or sentence structure, diction, in the sense of gross errors of word choice, punctuation, and spelling (The ability to use the conventions of writing: in basic form mechanics, in word-form mechanics)
 - C. Critical Thinking: Detection of unstated assumptions, perception of cause-and-effect relationships, and anticipation of the needs of the reader (The ability to express oneself logically in writing)

- D. Effectiveness: Adequacy of emphasis, adequacy of development, exactness of expression, economy, simplicity, and variety
- E. Appropriateness: Choice of level of usage suitable to purpose and reader; i.e., using the right "tone" and appropriate diction and employing tact where desirable

2. Cooperative English Tests (1960 Revision)

A. English Expression

- i. Effectiveness of Expression (sentence completions with words, or phrases, or clauses - choose the best one of four alternatives)
- ii. Mechanics (three line sentences, some with errors, some with no errors - testing grammar, mechanics, punctuation, spelling diction, effectiveness in sentences)

B. Reading Comprehension

- i. Vocabulary (choose the best one of four alternatives)
- ii. Reading (passages varied in style and content with questions ranging from recalling facts to interpretation)
 - a. Level of Comprehension
 - b. Speed of Comprehension

3. Writing Skills Test by Science Research Associates, Inc.

- A. Vocabulary (choose from among four expressions the one closest in meaning to a selected word)
- B. Sentence Recognition (indicate the number of complete sentences contained in each item)
- C. Grammar (indicate the part of a sentence - A, B, C, D, or E - containing a mistake in grammar)
- D. Punctuation and Mechanics (indicate the location of the error for each item)

- E. Spelling (find the misspelled word in a list of items and indicate how it should be corrected)
- F. Sentence Building (rewrite each sentence, starting with the word or words given, and then answer the question about the rebuilt sentence)

An examination of these three objective tests in relation to the concept of writing ability delineated in Chapter II of this study revealed to the investigator that the selection of the Cooperative English Tests and the STEP: Writing would provide, with a minimum of overlapping, broad testing coverage of a variety of writing skills. These tests are described further in a following section of this chapter.

Summary

From this review of objective testing of written composition several points are pertinent:

1. Early objective tests seemed to measure grammatical correctness as an objective of writing.
2. There is a growing tendency to attempt to design objective tests to measure a variety of abilities and skills needed for written composition; or, in other words, there is evidence that test-makers are becoming concerned with broader aspects of writing than correctness in usage, spelling, and vocabulary.
3. The more recent objective tests attempt to analyze the process of writing more thoroughly and attempt to measure more of the phases of writing.
4. A variety of objective exercises are in existence: inter-linear exercises, construction-shift exercises, error recognition, usage, sentence correction, paragraph organization, organization, conventions, critical thinking, effectiveness, and appropriateness.

5. Several objective tests have proved to be good predictors of writing performance when they are checked against a highly reliable criterion measure.
6. Many necessary skills and abilities are not yet tested objectively. (This point is further developed in Chapters III and IV.)
7. There is a necessity for a large-scale investigation to determine how several current standardized objective tests compare with each other, to determine what combinations of these tests account for the greater amount of variance of written composition scores, and to determine the amount of variance accounted for by all tests.

II. THE SELECTION OF STANDARDIZED TESTS FOR THIS INVESTIGATION

The three objective tests selected for inclusion in this investigation are described in this section of the present chapter. Justification for selecting these tests, the skills these tests purportedly measure, and the reported reliabilities and validities of the tests are also presented.

Sequential Tests of Educational Progress: Writing, Form 2B

Description. The publishers of STEP Writing, Form 2B claim that this test:

Measures ability to think critically in writing, to organize materials appropriate for a given purpose, to write effectively, and to observe conventional usage in punctuation and grammar. Materials were selected from actual student writing, in letters, answers to test questions, newspaper writing, announcements, essays, reports, records, minutes, logs, stories, notes, outlines, answers to questionnaires, and directions (Manual for Interpreting Scores: Writing, 1957, p. 7).

Though the "committee of specialists" responsible for the STEP Writing tests does not claim that the above five categories are either mutually exclusive or exhaustive, the test weights the five skills approximately as follows:

Organization.....	20%
Conventions.....	20%
Critical thinking.....	15%
Effectiveness.....	30%
Appropriateness.....	15%

The publishers of the STEP Writing test state that the "ability to identify errors or weaknesses in writing is not necessarily indicative of a student's ability to produce writing that is free of the same errors or weaknesses" (Manual, 1957, p. 9). For this reason the items included in the STEP Writing test are, in most instances, revision items; that is, the student must identify an error or weakness in a passage and he must select a revision which most satisfactorily solves the problem in relation to the purpose of writing. The test booklet contains 60 items of the multiple-choice type; the student has to choose the best answer from among four choices presented. The items are equally divided between two parts, so that a test may be given in one or two sessions, depending on class schedules. A single score based on "number right" is obtained.

Validity

With reference to the STEP Writing, the publishers state that content validity of the STEP Writing is of primary importance and that

content validity is best insured by relying on well-qualified persons in constructing the tests. However, as an empirical check on this armchair approach, the investigator is not aware of any validity studies relating test scores on STEP Writing 2B (or 2A) to suitable criterion measures. Writing in Buros, The Sixth Mental Measurements Yearbook, Allen states:

The burden of demonstrating usefulness falls on publishers. So it is remarkable that, among all the statistics relating to the test (STEP Writing) none can be found comparing STEP-W scores with any other measure of writing, including even the companion STEP Essay writing test. Five years after publication, STEP Writing is still an attractive test of uncertain worth (1965, p. 597).

Thus, with inadequate statistical evidence of validity and reliability, judgement about the worth of STEP Writing has to be made by the investigator in the light of the very detailed presentation of technical information, regarding development, norming, and statistical methods accompanying the test. Furthermore, the decision to include the STEP Writing has been made on the evidence that the test is carefully planned, and ambitiously conceived, and that it theoretically accounts for part of what the investigator considers aspects of the concept of writing ability and the writing process. The bases for the investigator's concept are discussed in Chapter II.

Reliability

Reliabilities reported for the STEP Writing are the results

of internal analyses based on single administrations of the tests. They are therefore estimates of internal consistency. Correlations between scores on alternate forms or between test-retest scores have not been obtained. The publishers in the Technical Report accompanying the STEP tests report a reliability of .85 for the STEP Writing 2A, with 486 grade eleven students (1957, p. 10).

Cooperative English Tests: Reading Comprehension, Form 2A

Description

The Cooperative English Reading Comprehension Test is divided into two parts: Part I is a Vocabulary Test and Part II, a Reading Comprehension Test. In Part I the student is asked to look at a word and to choose from a list of four words or phrases below it, the one which most nearly approximates the same meaning. The items range in difficulty from easy words with widely divergent meaning choices to difficult words with choices separated only very slightly in meaning. The raw score is simply the number of items right out of sixty.

Part II consists of a Reading Comprehension Test which is itself divided into two sections, one representing the student's level of Comprehension and the second his Speed of Comprehension. Though the test provides separate scores for the level and speed of comprehension, for the purpose of this study, only vocabulary, comprehension, and total scores were used. The reading passages are

varied in content, so that they represent material similar to what students read in school. The questions on each paragraph passage range from those requiring the student to recall a "fact" of the passage to increasingly complex questions requiring him to interpret what he has read. For each question the student chooses the best of four completions presented.

Justification for the Inclusion of a Reading Comprehension Test

It is an educational assumption that there is a link between reading and writing activities, though, admittedly, there is limited experimental evidence to validate this assumption. One notes that many "accomplished" writers have testified to the power of good books in their lives (Lowe, 1964; Cowley, Writers at Work: The Paris Review Interviews, 1958); but from the research point of view personal testimony is of limited value as evidence.

Though not exploring the nature of the reading-writing abilities of high school students, the work of Heys (1962), Christiansen (1965), Hinkelman (1956), and Jones (1966) offer somewhat limited support to the assumption. Heys concluded that: "For many students reading is a positive influence on writing ability" and "the influence of reading on the ability to write appears to be a separate factor not directly related to teacher's personality and enthusiasm" (1962, pp. 321-322). Following Heys' work, with a study of similar design, Christiansen concluded that "the reading done in

the control classes did as much to promote growth in writing as did the writing of sixteen extra themes in the experimental classes (1965, p. 131). Again, Hinkleman, in a study concerned with the role of reading ability in school achievement found the coefficients of .94 for the 2-A grade, .75 for the 5-A grade, and .82 for the 7-A grade between reading ability and written composition (1956, pp. 65-67). Jones, in a study designed to ascertain whether students' written composition ability (as measured by the STEP Essay Test, Form 2A) was related to students' reading ability (as measured by the Cooperative Reading Comprehension Test, Form 2A), found that "there is a significant relationship between students' written composition scores and their reading ability scores when the relationship is measured in the presence of the students' intelligence scores, the socioeconomic level of the family, the sex of the pupil, and the teacher variable" (1966, p. 82). However, when, in the Jones study, the total reading score was broken into its component parts of vocabulary scores and reading comprehension scores, it was found that students' vocabulary scores (in the presence of the other variables) accounted for a greater part of the variance of essay scores than did the reading comprehension scores.

Basically, justification for the inclusion in this investigation of an objective test on reading comprehension should be determined on the basis of an examination of the reading-writing relationships. Are reading about ideas and writing about them

somewhat reverse procedures? What skills do these procedures have in common? To what extent would written responses elicited by particular questions reflect the patterns of meaning found in the materials read by students?

As early as 1917 Thorndike described reading as

a very elaborate procedure involving a weighting of each of many elements in a sentence, their organization in the proper relations to one another, the selection of certain of their connotations and the rejection of others, and the cooperation of many forces to determine final response (1917, p. 323).

In an attempt to explore the "psychological nature of reading ability", Davis, using factorial analysis, posited nine basic skills as essential to reading comprehension. These were:

1. Knowledge or word meanings
2. Ability to select the appropriate meaning for a word or phrase in the light of its particular contextual setting
3. Ability to follow the organization of a passage
4. Ability to select the main thought of a passage
5. Ability to answer questions that are specifically answered in a passage
6. Ability to answer questions that are answered in a passage but not in words in which the question is asked
7. Ability to draw inferences from a passage about its content
8. Ability to recognize the literary devices used in a passage and to determine its tone and mood
9. Ability to determine a writer's purpose, intent, and point of view, i.e., to draw inferences about a writer (1944, pp. 185-197).

Other studies exploring the complexity of reading comprehension through the use of factor analysis have indicated what appears to be general language and thinking factors. Sochor (1959), in attempting to combine the factors reported by Davis, Langsman, Anderson, and Gens states:

If factors are interpreted and grouped, there seem to emerge...A "word factor" dealing primarily with individual word forms and their meanings appears to be one...A second might be called a "verbal factor" which contributes to the ability to see interrelationships among ideas represented by words in context but which would not involve too much abstract reasoning...A third factor "abstract reasoning" appears in all studies...Language, meaning and thinking continua seem to operate together among the factors...in an interrelated manner at all times in any given reader, each influencing the others in a constantly changing pattern (1959, pp. 4-5).

Reading comprehension is thus a thinking process, a process which requires mental facility in manipulating verbal concepts, a background of experience, and skill in the specialized techniques of reading comprehension, combined with skill in the mechanics of reading. For Lampard the basic reading skills for high school students are:

1. An ability to recognize and associate words and their appropriate meanings,
2. An ability to use context to obtain the meaning of words,
3. An ability to extract and differentiate the main ideas and the supporting detail in a passage,
4. An ability to make inferences and interpretations,
5. An ability to follow the organization of a passage and identify antecedents and referents,

6. An ability to determine the writer's purpose (1966, p. 45).

As outlined in Chapter II, effective writing can also be viewed as involving an active reasoning and an associational process; or in terms of Sochor's factors, it can be seen as involving a "word factor", a "verbal factor", and an "abstract reasoning factor". Embedded in the investigator's concept of writing ability and the writing process were the following writing skills:

1. An ability to use words correctly, a feeling for words, alertness to distinctions, implications, and overtones, and knowledge of apt, precise, and vigorous expression.
2. An ability to generate and express ideas with fluency.
3. An ability to see and make connections and interactions between the parts of the communication and to check hypotheses with proof. An ability to understand and recognize the basic kinds of relationships among particulars and between particulars and generals, and an understanding of the number and kind of particulars needed to develop selected topics.
4. An awareness of form, involving organization in the sentence, the paragraph and the essay as a whole. Introductions to prepare the reader, clear statement of central idea, development of ideas, transitions, and definite conclusions. An ability to formulate an outline.
5. An ability to write with purpose, intent, and point of view, a knowledge of the strategies of rhetoric, the requirements for securing the reader's attention, the dynamics of persuasion and the resources of style.

Though research studies have shown that there is a relationship between reading and writing activities, the nature of this relationship or the quality and structure of written responses to reading have seldom been explored. One may justifiably ask whether the above

listed writing skills can be termed reading-writing skills, and if so what elements are essential to this relationship. For example: (1) What is the relationship between the ability to recognize and associate words and their appropriate meanings with the ability to use words correctly in writing? (2) What is the relationship between the ability to follow the organization of a reading passage, and the ability to formulate an outline for writing? (3) What is the relationship between the ability to determine the writer's purpose when reading, and the ability to write with purpose, intent, and point of view? (4) What is the relationship between the ability to extract and differentiate the main ideas and supporting detail in a reading passage and the ability to see and make connections and interactions in writing?

Though lacking empirical evidence on these questions, it nevertheless appeared reasonable to the investigator that the inclusion of an objective test designed to measure some of the above reading skills would be appropriate for this investigation on the assumption that reading and writing skills have a common core of reading-writing elements.

Validity

The validity of a reading test is determined by the extent to which the test measures the skills actually involved in the reading process. The publishers claim that content validity is best insured by relying on "well-qualified people" to make a thorough analysis

of the reading process and then to construct items to measure the identified skills. This process was followed by the Cooperative Test Division in constructing the Cooperative English Reading Comprehension Test. It should be noted that the present Cooperative English Reading Comprehension Test was constructed (revised) in 1960 and thus very few studies are available which make specific reference to the test's validity. Flemming, in his review in The Sixth Mental Measurement Yearbook, points out that during the period Cooperative English Tests have been in use they have undergone substantial revisions and a great deal of background and interpretative information has been accumulated (1965, p. 806). In the Technical Report the publishers provide the results of eighteen studies concerned with the predictive validity of earlier forms of the Reading Comprehension Test. One such study reported is that by Traxler. Using teacher's ratings of reading ability as the criterion measure, Traxler found the predictive validity coefficients of the Reading Comprehension sub-tests to be as follows: Vocabulary, 0.57; Speed of Comprehension, 0.56; and Total Reading Comprehension, 0.61. These coefficients were obtained for 115 eleventh grade students. Also, as noted earlier, Jones using a criterion of actual student writing reported correlations of 0.52 for the vocabulary section of the test and a correlation of 0.47 for the reading comprehension section, for the 147 grade ten students.

Reliability

The publishers in the Technical Report accompanying the Cooperative English Reading Comprehension Test report the following reliability coefficients between parallel forms (Form 2A and Form 2B) of the Reading Comprehension Test with 780 grade ten students: Vocabulary, 0.89; Level of Comprehension; 0.78; Speed of Comprehension, 0.87; and Total Reading Comprehension, 0.94 (1960, p. 19).

Cooperative English Tests: English Expression

In a manual accompanying the English Expression test the publishers state that the test "is not a direct measure of writing ability, but evidence suggests that ability to do well on this kind of test is related to ability to write well in an 'essay' situation" (Manual for Interpreting Scores, 1960, p. 7). Writing in The Sixth Mental Measurements Yearbook, Sherwood reports of an experiment at the University of Florida which produced a correlation of 0.54 between Form 1A and essay examinations (1965, p. 558). Sherwood also indicates in his review that "the Cooperative Test more closely approximates the ordinary writing process than some" (other objective tests), (p. 558).

More specifically, the test of English Expression divides itself into two sections with the total English Expression score based on correct answers to Part I and Part II of the test.

"Effectiveness" is the title of Part I of the English Expression tests and refers to the choice of the written expression which precisely conveys the meaning intended. About two-thirds of the items in this part require the student to complete a sentence by choosing the best single word from among four presented. The remaining items require completion of sentences with the most effective phrases or clauses.

"Mechanics" is the title of Part II of this test and refers to usage, spelling, punctuation, and capitalization. A new item type is introduced here: the student is given a three-line sentence and asked to identify the line in which any mechanics error occurs or, alternatively, to indicate that the sentence contains no error.

The new item type was developed to simulate the "real" situation in which a student goes over his own or another's writing to find errors. (Cooperative English Tests: Manual for Interpreting Scores, 1960, p. 6).

On the basis of the above description of the English Expression test, it is indeed obvious that the test does not measure the ability to write, but it certainly does attempt to measure selected critical powers which the researcher has suggested are related to the ability to write. Students are required to make rational choices among words. The test also has relatively good coverage of "expression", of some details of composition - sentence structure, diction and the like - as distinguished from such broader aspects as unity and organization. Furthermore, in one way, the mechanics section of this test simulates the period of proofreading and revision - one part of the writing process.

Summary

The selection of these three standardized tests provided objective testing on several germane skills and abilities, all of which are incorporated into the explication of writing ability and the writing process presented in Chapter II. Though overlapping exists on the testing of some particular skills, the following are listed:

Sequential Tests of Educational Progress: Writing

Organization: Reasonable ordering of ideas, events, facts, etc. (The ability to organize materials: in the whole passage, in the paragraph, and in the sentence)

Conventions: Attention to syntax or sentence structure, diction, in the sense of gross errors of word choice, punctuation, and spelling

Critical Thinking: Detection of unstated assumptions, perception of cause-and effect-relationships, and anticipation of the needs of the reader

Effectiveness: Adequacy of emphasis, adequacy of development, exactness of expression, economy, simplicity, and variety

Appropriateness: Choice of level of usage suitable to purpose and reader; i.e., using the right "tone" and appropriate diction and employing tact where desirable

Cooperative English Test: Reading Comprehension

Vocabulary: Recognition vocabulary: items using words of increasing difficulty and items requiring increasingly close discrimination among the choices

Comprehension: Passages varied in style and content with questions ranging from recalling facts to interpretation

- determining the meanings of words from contextual clues
- organizing meanings
- construing the writer's meaning
- drawing conclusions from the content

Cooperating English Test: English Expression

Effectiveness: Sentence completions with words, or phrases, or clauses - choose the best one of four alternatives

Mechanics: Three line sentences, some with errors, some with no errors - testing grammar, mechanics, punctuation, spelling, diction, effectiveness in sentences

III. THE CONSTRUCTION OF OBJECTIVE TESTS FOR THIS INVESTIGATION

The Written Expressional Vocabulary Test and The Sentence Sensitivity Test, the two tests constructed by the investigator, are described in this section. Specifically, the justification for the tests, the description of the tests, the procedures followed in constructing the tests, and the results of a pilot study are presented.

Written Expressional Vocabulary Test

Justification for the Test. A dictionary may define vocabulary as "all the words of a language" or "all of those words used by a particular person, group of persons," or, often, "a list of words usually arranged alphabetically and defined." On the surface, it appears that the term vocabulary may be roughly equated with a stock of words. - Researchers, though, refer to speaking vocabulary, writing vocabulary, listening vocabulary, and reading vocabulary. Such divisions are made on the basis of the "easily verifiable evidence that a person may understand a word he hears and yet not use it in his writing or speaking. Too, he may read a word and be able to determine its meaning and yet not be able to do equally well with the same word in a listening situation" (Petty, Herold, and Stoll, 1968, p. 10).

In this test, however, the investigator was primarily concerned with the words that belong to a student's written expressional vocabulary - with the words known well enough to be used in writing with sufficient cognizance of their major variations in meaning. It was the investigator's aim to construct an objective measure of students' written expressional vocabulary and to use this measure as an index of students' skill in written composition.

As elaborated in the discussion of writing ability and the writing process in Chapter II, the value of the correct choice of words (in accordance with the outlined semantic relationships) is of prime importance for the student in his attempt to adequately express the ideas he wants to communicate. Thus, a first consideration in constructing an objective test to measure students' written expressional vocabulary was to determine the main dimensions to be measured. Three dimensions were identified: variety in expression, precision in expression (exactness, effectiveness), and an appropriate level of usage.

The desirability of varied and precise expression is a point well-emphasized in written composition textbooks. In Composition for Canadian Universities, Waterston and Beattie state:

If the communication is to be complete and effective, words must be used with precision. For the communication of complex ideas an extensive and varied vocabulary is indispensable...The would-be writer should cultivate with special energy his feeling for words...No one can doubt that an extensive vocabulary is essential for success as a writer. A wider, more varied, and more precise vocabulary increases a student's capacity to think clearly, to differentiate exactly between the many ideas presented to him (1964, pp. 165-166).

Similarly, Corder, in Rhetoric: A Text Reader on Language and Its Uses, states:

You must say precisely what you mean; coming near to it, as you will discover, is not enough, unless of course, you are content with coming close. But to say precisely what you mean is an arduous task. It is entirely unlikely that the right word will appear magically upon your page unless you have created the magic with deliberation. For there can be no other mode: deliberation must be your manner.

The words you use in writing must, to be sure, be correct and natural...But correctness and naturalness, while they may be conditions of good writing are not sufficient alone to make good writing. Preciseness is far more crucial in your choice of vocabulary (1965, p. 220).

Variety in expression and precision in expression are undoubtedly two of the main dimensions in students' written expressional vocabulary. But anyone who is conscious of his use of language knows how he diversifies it according to circumstances and occasion. Usage must be appropriate to the context. Accordingly, the student makes adjustments in sentence structure, in rhythm, and in diction as occasion seems to demand. "Choice of words depends, therefore, upon a well-developed sense of 'levels of usage' and upon the ability of the writer to frame his thoughts suitable to each occasion" (Waterston and Beattie, 1967, p. 177).

Composition textbooks also provide a consensus on "specific principles" or "characteristic ways" in which words are used to achieve the three dimensions of variety, precision, and appropriateness. For Cedric Gale (1966, pp. 230-235) words, to be effective, should be selected according to twelve "principles"; for Waterston

and Beattie (1964, pp. 172-177) there are nine "methods" of imparting vitality to written style; Hamalian and Volpe (1960, pp. 256-263) suggest five "main considerations" to assist the writer in determining his choice of words; while Corder (1965, pp. 220-243) exemplifies, with many brief illustrations, four "characteristic ways" in which a writer can obtain precision. It should, however, be noted, that in all the above statements, the authors suggest that the "principles" or "characteristics" should, like all rules about writing, be interpreted with some sense of imagination and should not be followed so slavishly as to produce affectation.

Representative of points from the four composition textbooks listed above are the following "specific principles" or characteristic ways" in which words are used by students to achieve variety and precision in their writing:

1. through the use of concrete and abstract words;
2. through the use of specific and general words;
3. through the denotative and connotative use of words;
4. through the use of synonyms; and
5. the avoidance of vague words, overworked words, trite phrases, ambiguity, redundancy, pretentious words, cliches, incongruity, vulgarisms, improprieties, and euphemisms.

These five "principles" reflect the students' problem of choosing words; that element of choice which the investigator placed as central in his explication of the writing process and the concept of writing ability.

To provide a further rationale for the types of exercises included in the Written Expressional Vocabulary Test a brief explanation of these five principles is presented. The "good writer" communicates his meaning by judicious choice and combination of abstract and concrete language. He selects his words in accordance with the demands of the specific situation and with the type of reality to which he refers. It is the writer's task to find the word so fitting that the meaning the writer puts into it and the meaning the reader takes from it coincide. The writer who is the most successful in doing this is likely to be the writer who most consciously and consistently uses concrete words. Yet our language contains many abstract words, and no one can safely contend that they do not play an important part in effective writing. In some theoretical discussions, in some evaluations and exhortations logically developing traditional knowledge, in some exercises in moral judgement, abstract words can be strikingly exact. However, abstract words are somewhat ineffective when employed where concrete words would work better. Similarly, abstract words can be vague if they are used too often and too exclusively and are not supported, clarified, illustrated, or defined with concrete words.

Too, writing that has a preponderance of general words is likely to seem heavy and pretentious. However, general terms are as necessary to the vocabulary as specific terms. When, for example, the subject is philosophical - whenever, indeed, the subject calls for generalizations - the writing will have a large proportion of general

terms; but in descriptive and narrative writing the tendency should be toward the specific. The ideal is a judicious balance. The general terms must have enough specific terms to support them; but specific terms should not be used so freely and exclusively that the reader is confused by sheer detail.

The careful stylist also observes, as another principle of word usage, the denotative and connotative use of words. To communicate his ideas so that others can understand them, a writer needs to have a sure grasp of the denotative meanings of the words he uses. But to make the choices that lead to responsible, effective expression, he needs also to be sensitive to the connotation of words.

Connotative words, often referred to as emotive, evaluational, intentional, loaded, or slanted words, give writing a quality that can be called suggestion. The connotation of a word expresses more attitude than content, and one must choose words carefully if he wishes to control the attitude he expresses. The writer cannot ignore the suggestive power words have, a power they acquire through the way in which they are commonly used or by association with a particular context. In reacting intelligently the writer must separate the connotative from the denotative meaning of each word, and, he must ask himself whether the actual situation in which the word is used justifies the connotative term. The careful stylist is both gratified by the abundance of words at his disposal and challenged by the need to recognize shades of implication.

The use of synonyms exemplifies a further challenge for the writer. In the preparation of a theme, the selection of synonyms is often a matter of substituting for a word that is vague, colorless, or inappropriate a word that is exactly right; a word that fits the context - in meaning, in connotation, and in tone.

Finally, the careful stylist avoids vague words, over-worked words, trite phrases, ambiguity, redundancy, pretentious words, cliches, incongruity, unless such words are used to achieve an intended effect.

Essentially, then, it was the investigator's intention to obtain an estimate of students' written expressional vocabulary along three dimensions: variety in expression, precision in expression, and appropriateness of usage. More specifically, the investigator sought to test students' knowledge of these three dimensions through the use of questions based on the suggested five "specific principles" or "characteristic ways" in which words are chosen and to use this estimate as an index of students' skill in written composition.

Description of the Test. In measuring variety of expression (or range in expressional vocabulary) the investigator attempted to ascertain to what extent students, when they are presented with a particular concept, are able to express it through the use of different words. In measuring precision of expression the investigator sought

to determine the degree of discrimination which students exercise in selecting one of a number of synonymous or nearly synonymous terms as most appropriate to represent the nuance of meaning intended to be conveyed in a particular context.

Though it was somewhat easy to distinguish theoretically among the dimensions of variety, precision, and appropriate usage it was difficult to formulate questions to test the dimensions separately. Thus, one can observe in the test, as constructed, some questions which test both the dimensions of variety and precision and one can observe questions which overlap on some of the five "characteristic ways" words can be used to achieve variety and precision. However, it was the students' total score on the Written Expressional Vocabulary which was used as a predictor of students' written composition performance.

A copy of the Written Expressional Vocabulary Test is included in Appendix B. The nine subtests of objective (and semi-objective) questions used are described as follows:

Subtests (1) and (2), and to some extent subtest (3), were based on the synonym level of response. As in their own writing, students were in these situations faced with the problem of finding words to express a concept and of striving for different words to express nearly the same thought.

Example from Subtest (1):
to predict the future _____

Example from Subtest (2):
He was captured by the wild tribes. _____

In addition to knowledge of synonyms, Subtest (3) tested the students' ability to make action and movement highly specific, that is, the ability to replace general words with words to flash a vivid picture on the minds of the readers.

Example from Subtest (3):

"No," the girl _____, and left the room."

Suppose "the girl" left speaking loudly and displaying angry passion or vehemence. _____

Subtest (4) consisted of two-word completion exercises designed to test students' precision and discrimination, and to some extent the students' ability to comprehend through clue words the meaning of a sentence still incomplete.

Example from Subtest (4):

A _____ response is one that is made with _____.
 (a) stupid, fear (b) speedy, alacrity (c) sure, slowness
 (d) harmful, grimaces (e) pleasant, surmise

Subtest (5), a one-word completion test, focused on precision, discrimination, and denotation. Students were to select the correct answer from the four words suggested.

Example from Subtest (5):

I should not feel confident in venturing on a journey in a foreign country without a(n) _____.
 (a) associate (b) chum (c) companion (d) friend

In Subtest (6) the students were to associate the individual words from a prepared list with the appropriate character drawings.

Subtest (7) required the students to supply one word that could be substituted for the parenthetical expression embedded in each sentence.

Example of Subtest (7):

I was a very (small and thin) lad. _____

In Subtest (8) the students were required to select from word groups the word having the most favorable connotation.

Example from Subtest (8):

(a) dictator (b) strong man (c) absolute ruler _____

Finally, in Subtest (9) the students were required to replace cliches, etc., with more appropriate words.

Example from Subtest (9):

I would give you some violets but they are (dead as a doornail). _____

Procedure for Construction

The procedure used in constructing the Written Expressional Vocabulary Test followed somewhat closely the general steps in building any achievement test. The dimensions, as described above, which were thought to be germane to the students' struggle for the effective use of words in written composition were first identified. The decision of how the investigator was to measure these dimensions was made after considerable study of many present vocabulary tests, after reading numerous books on vocabulary development and written composition, and after discussions with several high school English teachers and with colleagues at the university. Largely, too, the decision was made on the basis of the investigator's knowledge of the vocabulary development and "effective" writing of secondary school students. After a sufficiently large pool of items was collected and organized into nine different subtests, the items were then reviewed by colleagues and the investigator's advisor. Items were subsequently revised by the investigator. The preliminary Written Expressional Vocabulary Test was then administered to a class of grade ten students who were not to be part of the investigator's final sample.¹ An item analysis was conducted and a revision

¹The results of this administration are found in the Pilot Study described at the end of this chapter.

of the test made before it was used in this investigation.

Sentence Sensitivity Test

Sentence Structure and Prose Quality. The decision to construct an objective (or semi-objective) test on sentence sensitivity and to include such a test as one of the predictor variables is based on the assumption that a significant relationship exists between the general merit of tenth grade writing and certain aspects of sentence structure and sentence sensitivity. There is the further assumption that the ability to identify well-formed, mature, and emphatic sentences is related to students' skill in using such sentences in their writing. However, the test constructed requires both the identification of such sentences and the composing of other sentences.

Today, the researcher who turns to professional literature for assistance on the grammar-composition problem (or, more specifically, sentence structure and prose quality) quickly becomes aware of a wide variety of findings. One first notes that practically all studies prior to Loban's study in 1963 followed an almost standard procedure by determining mean sentence length, mean clause length, and subordination ratio. Among those investigators using what one might call methods of traditional grammar are Bear, Davis, LaBrant, and Watt. LaBrant (1933), using the subordination index (the frequency of subordinate clauses expressed as a percentage of all clauses), studied the language development of children in grades four through twelve, and found an increasingly larger number of subordinate

clauses used, with a correlation between chronological age and subordination of .41. Watt's (1944) study of the mental development and development of written expression in English children produced similar findings, with the percentages of dependent clauses at the various ages being as follows:

Age	7.8	8.9	9.10	10.11	11.12	12.13	13.14	14.15
%	16.2	22.1	27.40	32.30	36.30	39.50	41.30	43.20

Frogner in a well-known 1939 study tested a functional grammar approach to composition instruction against an approach that ignored grammatical terminology and concentrated on the grace and effectiveness of the communication of the writer's thoughts. She found that although the "grammar approach" produced some improvement, the "thought approach" taught students more skills in less time. Writing skills gained through the "thought approach" stayed with the students longer than did those gained through the "grammar approach" (1939, p. 525).

More recently, investigators such as Loban, Hunt, Armstrong, Bateman and Zidonis have, in their search for indices of language development, been somewhat successful in using the methods of modern linguistics. From the work of Loban a most significant feature to emerge is that:

Not pattern but what is done to achieve flexibility within the pattern proves to be a measure of effectiveness and control of language at this level of language development...In the movable elements of the patterns, the high group consistently shows a greater repertoire of clauses and multiples (movables within movables)...For subject nominals, the low group depends almost exclusively on nouns and pronouns. The high group can use noun clauses, infinitives, and verbals (1963, p. 84).

Though using a transformational description of sentence structure, Hunt designed a study to identify developmental trends in the frequency of various grammatical structures written by students of average intelligence in the fourth, eighth, and twelfth grades. Using a 1,000 word sample of writing for each of nine boys and girls in each of the three grades Hunt analyzed (in addition to sentence length, clause length, subordination ratio, and kinds of subordinate clauses) the following structures within a clause: coordinated structures, nominals, the verb auxiliary, main verbs and complements, modifiers of verbs, and predicate adjectives. Though the usefulness of Hunt's findings may be somewhat limited by his small sample and his not distinguishing between well-formed and mal-formed sentences, his study does point to what may be a most significant index of maturity in language.

T-unit length is somehow tied closely to maturity (in writing)....It is the absorption of kernel sentences that appears to account for the lengthening of T-units, both single clause and multiple clause.... Younger students tend to use short clauses to express these meanings, whereas older students tend to reduce such clauses to mere modifiers which are consolidated with the same noun in another clause, thus achieving greater length....The chief factor which lengthens clauses appears to be the increasing of the noun clause modifiers of nouns and the nominalization of the clauses. This factor and the increase in adjective clauses account in the main for the increased length of T-units (1965, pp. 141-143).

Finally, Potter, in a comparative analysis of twenty samples of "good" and "poor" tenth grade writing, questioned the usefulness of employing some of the above measures of maturity as measures of quality and suggested several other ways of measuring quality of writing. His study was designed to explore some of the grammatical differences between good and bad writing. The two groups of papers were examined first with regard to several established criteria that had been found by other studies to differentiate between "mature" and "immature" writing. These included LaBrant's subordination index, Hunt's minimal terminable unit or T-unit length, and sentence length. Potter states:

Because of the similarity in the two groups of papers of many gross measurements (the subordination index, for instance, as well as the sentence length/T-unit length ratio, was found to be exactly the same), it seems likely that future investigators will find more meaning in the way particular structures are used than they will find in gross totals (1967, p. 27).

The following statements by Potter, on the basis of his research, suggest these profiles of the good and poor writers:

The poor student writer tends to use more of his words in simple basic sentence patterns and correspondingly fewer in internal structures of modification.

The good writers in this study were much more proficient than the poor writers in the use of coordinating words and devices used to add T-units to grammatically complete sentences. The good writers made much more use of the conjunction for and the semicolon.

Apparently, what many poor writers need is instruction in the judicious use of the passive voice, not warnings against a structure they seldom employ in any fashion.

The poor writers, however, used twice as many conditional clauses as the good writers used.

The good writers proved superior in modifying their objects with prepositional phrases, clauses, and verbal structures, as well as in using clauses and verbal structures themselves as objects.

The good writers among the subjects of the study tended to use the optional that to introduce nominal clauses, while the poor writers tended to ignore this somewhat literary device.

The good writers did, however, appear to have a greater command of the many clause introducing conjunctions.

The good papers contained more verbal structures of every type than did poor papers (1967, pp. 17-28).

Since the above research studies indicate that measurable differences in sentence structure can be found between the "good" and "poor" writing of average students, such evidence should

indeed be taken into consideration in constructing any objective test to measure students' sentence sensitivity as a skill basic to "good" written composition. In addition to research studies, some recent professional articles as well as grammar and composition textbooks emphasize the importance for the student to acquire the methods and techniques of sentence building and to achieve variety within basic sentence patterns.

A representative text emphasizing the importance of the sentence in the study of written composition is Structural Grammar and Composition by Hulon Willis. A main thesis in Willis' book is that "the sentence is in a true sense a composition in miniature," and that the individual sentence is one of the principal keys to good writing.

No matter how extensive a piece of writing is, its quality finally rests in the quality of its individual sentences. If a person can write well-formed sentences - clear, precise, mature sentences - he can quickly learn (or, more likely, already knows) how to develop larger units of writing. A person who is inadequate to the task of composing good sentences is already defeated in his efforts to produce a whole paper of quality. Of all the main principles of composition - organization of the whole paper, paragraph development, sentence composition, word choice, and correct usage - students can most fruitfully study the art of sentence composition. Learn to compose good sentences and all these other skills will naturally follow (1967, p. 19).

That the student who studies the art of sentence composition studies all the principles of composition is also a point well stated by Jackson Burgess:

All the basic principles of composition operate in the composition of the sentence. A good sentence, like a good essay or a good book (and far more than "a good theme"), requires a reasoned organization, a point of view, a consistent and appropriate tone, form, and diction. The principles of composition can be far more easily (taught) in the sentence than in longer compositions (1963, p. 259).

Conlin in Grammar for Written English (1961) and Newsome in "Expansions and Transformations to Improve Sentences" (1964) also emphasize the process of elaboration, that is, of nouns and verbs serving as nuclei or headwords to which modifiers are added to develop the meaning of an utterance. A variety of modifying structures gives language flexibility and makes possible to communicate a vast range of ideas with considerable finesse.

Stressing the importance of the sentence and somewhat bridging the gap between the emphasis on sentence structure by the traditionalists and that of the transformationalists, Francis Christensen states that "the most lucky find, the most radical insight to emerge from this inductive study, prompts the suggestion that our faith in the subordinate clause and the complex sentence is misplaced, that we should concentrate instead on the sentence modifiers, or free modifiers" (1967, p. xiii). Although granting that the balanced sentence deserves some attention in discursive writing, and the enormous range of coordinate structures deserves a bit more, for Christensen, however, "the rhythm of good modern prose comes about equally from the multiple-tracking of coordinate constructions and the down shifting and backtracking of free

modifiers" (1967, p. 2). Free modifiers give the skilled writer the options that rhetoric demands:

Bound modifiers are word modifiers, they are close or limiting or restrictive modifiers. They are in a sense obligatory, and, being obligatory, they do not give the writer the freedom of choice that rhetoric demands. Free modifiers on the other hand, are modifiers not of words but of constructions, from which they are set off by junctures or punctuation. Grammatically, they are loose or additive or nonessential or nonrestrictive. The constructions used are prepositional phrases or clusters; noun, verb, adjective and adverbial phrases or clusters; and one of the most important, verbid clauses or absolutes (1968, p. 577).

In addition to the above mentioned studies stressing the importance of sentence building and expansion through coordination, modification and subordination, other sources (particularly, high school composition textbooks) on effective sentence construction emphasize the importance of economy, emphasis, and unity. These concepts are also included in the investigator's Sentence Sensitivity test.

Description of Sentence Sensitivity Test

Seven types of exercises (subtests), testing different aspects of sentence making, are included in the Sentence Sensitivity Test.

Subtest 1. Students were to select from two ways of presenting an idea the sentence which they felt gave the greater emphasis to the idea the writer wanted expressed. This exercise tests more than clarity of expression. Specifically, the exercise

tests students' knowledge of some of the following ways of directing attention: the use of strong and effective words, the position or ordering of words, effective punctuation, intensives, repetition, inversions, climax, appositives, balance, passive and active voice, and effective use of imagery.

Subtest 2. This exercise of economy or conciseness required students to delete from selected sentences words not necessary for the sentence to convey full meaning. It tests students' ability to express thought as simply and as economically as the thought allows or requires.

Subtests 3 and 4. These tests required students to combine short sentences into one well-formed sentence. Subtest 3 required the combining of pairs of sentences into one sentence; Subtest 4, the combining of four short sentences into one sentence. These exercises measure students' flexibility and fluency with language resources in combining given facts and ideas into single, economical, and effective sentences. Generally, the exercises required that the students in writing the new sentence use some of the following principles: modification (of subject, of verb, of other sentence parts); coordination (of compound constituents such as compound subjects, predicates, objects, objects of preposition, objects of a verbal, and compounded word modifiers; and of compounded sentences); and subordination (appositives,

adjectival clauses, adjectival phrases, participial phrases, adverbial clauses and adverbial phrases).

Subtest 5. This exercise presented the students with a "first" sentence followed by several alternative "next" sentences from which they were to select the best "next" sentence. Students were to select the best "next" sentence on the basis of which most advances the thought, maintains suspense, uses words economically, fits the rhythm, maintains the mood, and avoids extraneous matter.

Subtest 6. This exercise required students to select from groups of four sentences the better formed, more mature sentence.

Subtest 7. This exercise, based on Christensen's work and examples, required the students to recognize the basic statement in the three statements presented and then to decide on the appropriate arrangement for the three statements. The exercise tests students' ability to write sentences rich in meaning by first deciding on the basic statements and then adding layers of meaning with word groups that make each basic statement less general, more specific and precise.

In constructing the Sentence Sensitivity Test the investigator assumed that questions representative of the above descriptions would provide an index of the students' ability in some aspects of sentence construction and that such an index would be predictive of the students' written composition performance.

Procedure for Construction

First, the investigator read widely in language development studies, studies of prose structure and written composition, and several high school written composition textbooks. The dimensions of "emphasis", "economy" or "conciseness", "layers of meaning", "best next sentence", "recognizing mature sentences", and "sentence combining" through modification, coordination, and subordination were selected as germane aspects of sentence construction for effective composition. Test items on these dimensions were reviewed by colleagues and the investigator's advisor, and subsequently revised by the investigator. The test was then administered to a class of grade ten students who were not part of the final sample. After an item analysis was conducted a further revision of the test was made before it was used as a variable in this investigation.

Pilot Study

In February, 1969, the two tests (the Written Expressional Vocabulary test with 66 items, and the Sentence Sensitivity test with 57 items) were administered as a pilot study to twenty-six grade ten students in a large composite high school in Edmonton, Alberta. The teacher's estimate of the students' scores in written composition (expository writing) was used as a criterion measure against which the two tests were validated. The data were analyzed, and the description of the tests which resulted is presented in Tables I and I'.

Tables I and I' show that the two correlations, 0.59 for the Written Expressional Vocabulary test and 0.50 for the Sentence Sensitivity test, with the teacher's estimate of students' scores in written composition are both significant at the .01 level. In view of these validity coefficients and the fact that the criterion was the teacher's estimate of students' scores in written composition it appears that the two tests are composed of reasonably valid items. The correlations of the individual subtests with the composition scores, along with the ranges, means, and standard deviations for the subtests are also shown in Tables I and I'.

Two aspects of item analysis were also examined: item difficulty and index of discrimination. Item difficulty is defined as the percentage of students who marked the item correctly. An index of discrimination is a correlation coefficient between the responses to a single item and total test scores. A high index indicates high item discrimination, as those who scored high on the test answered the item correctly and those who scored low answered it incorrectly. Summaries of the individual item statistics are shown for the different subtests of the Written Expressional Vocabulary test and the Sentence Sensitivity test in Tables I and I'.

In this investigation, indices of discrimination were calculated by dividing the scores at the median, with the top half forming the high group and the bottom half forming the low group. Thus, an index of discrimination was calculated as follows:

TABLE I
PILOT STUDY DATA ON WRITTEN EXPRESSIVE VOCABULARY TEST

Subtests	Corr. with Written Comp.	Corr. with Written Exp. Voc.	Max. Range	Range	Mean	Stand. Dev.	Level of Diff.	Index of Discr.
1	0.43	0.74	0-15	5-15	9.27	2.82	62%	.20
2	0.33	0.65	0-10	3-10	6.42	1.90	64%	.16
3	0.46	0.76	0-15	4-15	8.92	3.37	59%	.28
4	0.26	0.07	0- 7	0- 5	2.46	1.24	35%	-.01
5	0.01	0.07	0- 5	0- 4	2.46	1.10	48%	.09
6	0.51	0.74	0-20	3-18	10.92	4.01	55%	.29
7	0.31	0.33	0- 5	1- 4	2.89	0.82	55%	.06
8	0.35	0.58	0- 9	0- 9	4.00	2.04	44%	.21
9	0.07*	0.57	0-10	0- 8	5.62	2.06	56%	.17
total	0.59		0-96	34-76	52.96	11.63		

TABLE I¹
PILOT STUDY DATA ON SENTENCE SENSITIVITY TEST

Subtests	Corr. with Written Comp.	Corr. with Sent. Sens.	Max. Range	Range	Mean	Stand. Dev.	Level of Diff.	Index of Discr.
1	0.40	0.59	0-10	3- 9	6.54	1.86	65%	.23
2	0.08	0.59	0-15	2-14	9.12	2.72	61%	.18
3	0.45	0.76	0-20	6-17	11.96	4.08	60%	.29
4	0.29	0.36	0-10	3- 8	5.15	1.26	52%	.05
5	0.22	0.33	0- 5	0- 4	1.85	0.93	37%	.09
6	0.40	0.54	0- 7	1- 5	3.31	1.23	48%	.18
7	-0.11*	0.09	0-10	4- 9	7.39	1.39	74%	.00
total	0.50		0-77	33-56	45.31	7.35		

*With N = 26 a correlation of 0.50 is needed for significance at .01 level.

$$\text{Index of Discrimination} = \frac{H - L}{\text{the number of students in each half}}$$

The decision to accept or reject the individual test items was made on the basis of an examination of both the item difficulty and the index of discrimination for each item. Generally, with items in the middle range of difficulty (25% - 75%) a H - L difference was accepted as being significant if it was equal to at least 10% of the number of students in the sample. For both tests this significant H - L difference was 3 or an index of at least .20. When items were beyond this range of difficulty, that is when items were easy or difficult, a H - L difference of 5% of the sample was taken as acceptable.

Using these criteria as the main basis of judgement the following changes were made:

In the Written Expressional Vocabulary Test

- Subtest 4: two questions omitted
- 5: one question omitted, two questions added
- 7: two questions omitted, one question added
- 8: one question omitted
- 9: one question omitted, one question added

In the Sentence Sensitivity Test

- Subtest 1: one question omitted, one question added
- 2: three questions omitted, two questions added
- 3: two questions omitted, two questions added
- 4: the whole subtest omitted and replaced by questions testing other characteristics of sentence sensitivity
- 6: two questions omitted
- 7: four questions omitted, one question added

The final forms of the Written Expressional Vocabulary Test and the Sentence Sensitivity Test as used in this investigation are included in Appendix B.

CHAPTER IV

THE CRITERION MEASURE

I. VARIABLES ASSOCIATED WITH ESSAY TESTING

One of the fundamental problems in research into written composition is the general evaluation of actual writing.

Variables such as the assignment variable, the writer variable, and the rater variable have to be considered.

The Assignment Variable

Research studies concerned with the nature of the assigned essay topic inevitably point to the importance of selecting topics with care. It is a well-documented fact that if several topics are assigned as alternative topics from which one or two could be chosen by the student a student's rating might depend more on the topic he chose than on how well he wrote (Godshalk, Swineford, and Coffman, 1966, p. 13). To control the effects of the topic on the quality of writing a single topic should be chosen. However, when a researcher decides to eliminate choice and assign a single topic it becomes obvious that the topic selected should be familiar and of interest to most of the students in the group being tested. Diederich suggests that the topic assigned must be within the student's comprehension because:

Even the better students write badly when the topic is beyond them. Their struggles are apparent, not only in lack of organization and in vagueness of statement, but in the very structure of their sentences. Their knowledge of grammar, punctuation, etc., seems to desert them when they are labouring with an idea which is too big for them (1946, p. 585).

Diederich further maintains that the validity of an assigned essay topic can be questioned only if:

The students are required to write a paper extremely unlike the writing they will have to do in the normal course of events or if the papers are marked chiefly for their content; that is for the accuracy and completeness of the writers' knowledge of the subject assigned. If these two mistakes are avoided, the essay is unquestionably a valid test of ability to write, for it is an instance, a sample, of the very ability that one is attempting to measure. There is no more direct evidence of ability to write (1946, p. 584).

Procedures to Minimize the Assignment Variable. To offset the effects of the assignment variable the following procedures and steps were followed in this investigation:

1. No alternative topics or choices were given, thus controlling the effects of the topic on the quality of writing. All students wrote on each of three essay topics.
2. The types of writing responses the essay topics required were similar to the goals of composition instruction for the tenth grade, generally, expository-argumentative writing. For most students the type of descriptive writing response required by the twenty-minute paragraph was a subject of instruction in the previous year in grade nine.

3. An effort was made to select topics that would be within the students' comprehension and writing ability. Essay One (STEP Essay, Form 2A) was based upon a situation requiring analysis and some decision requiring interpretation, point of view, or a judgement to be stated and supported. Essay Two (STEP Essay, Form 2D) required students to write a character analysis based upon a speech given by a student to a school assembly. Furthermore, for Essays One and Two, the formulators of the STEP Essay tests, relying on knowledge of "expert teachers and supervisors of English", maintain that the topics were chosen so as to take into account socioeconomic, regional, cultural, and religious differences (Handbook for Essay Test: Level 2, 1957, p. 5). The third essay (a twenty-minute paragraph), "An Edmonton Morning at 20 Degrees below Zero," as selected by the investigator and three high school English teachers was thought to be appropriate since Edmonton's temperature was minus 20 degrees on the day students wrote the paragraph.

The Writer Variable

Many researchers purporting to measure writing ability actually only measure a student's performance on a given topic on a given day. Braddock, et. al. point out that composition examinations, although they are often referred to as measures of writing "ability", are "always measures of writing 'performance'; that is, when one evaluates

an example of a student's writing, he cannot be sure that the student is fully using his ability, is writing as well as he can" (1963, p. 6). The student is subject to a broad but finite range of distractions: he may suffer from personal concerns or from annoyances within the examination room.

Kincaid, in his study, "Some Factors Affecting Variables in the Quality of Students' Writing," concluded "that a single paper written by a student on a given topic cannot be considered as a valid basis for evaluating his achievement in a writing course at any time, unless that student's writing ability was rather low; and, even then a single paper would not provide an infallible basis for such an evaluation" (1963, p. 92). Kincaid also found that writing performance--especially the performance of the better writers--varies from day to day. Similarly, Anderson found that seventy-one percent of the fifty-five eighth grade students he examined on eight different occasions "showed evidence of composition fluctuation" apart from the discrepancies attributable to the raters (1960, pp. 95-96). Again, Diederich wrote in 1946 that about one fourth of a group of University of Chicago students changed their marks as a result of writing a second test but that less than five percent changed their marks as a result of writing a third. These studies of Kincaid, Anderson, and Diederich point clearly to the existence of a writer variable.

Further evidence of a writer variable is evident in a study by Traxler and Anderson who reported a correlation of .60 between two essays on almost identical topics, which were written within two days. The first essay was on "The discovery of Gold in California"; the second on "The Pony Express." The students were given an outline and unorganized notes for the writing of each paper and were told not to include any material that was not given in the notes. Though the papers were marked by two readers whose marks for the first essay correlated .94 and for the second essay .85, the correlation between the students' marks on the first papers and their marks on the second papers was only .60. This correlation indicates variation in the quality of writing, or the inability of students to perform at a consistent level on papers which were almost identical in form and content and had been written within two days (Traxler and Anderson, 1935, pp. 534-539).

To make allowance for this writer variable, Braddock states that "if it is desirable to evaluate a student's composition performance when it is as good as his performance typically gets, he should write at least twice, once on at least two different occasions, the rating of the better paper being used as a measure of his writing performance" (1963, p. 6).

Procedures to Minimize the Writer Variable. To reduce the effects of the writer variable students were required to write three essays on three separate days, with each essay being written on a different topic. The total amount of writing required was that which

students could do in 100 minutes -- two forty minute essays and one twenty-minute paragraph. The students' total score on the three essays was used as a measure of writing performance.

The Rater Variable

That inter-rater variation in composition marking exists is a well-documented fact. Investigations summarized by Ross and Stanley (1954) point out that beginning about 1900, several studies revealed that the marks received by students on compositions were more often a function of the personality of the rater than of the performance of the student. Lyman, in his Summary (1929), adds further support to the fact that disagreement among theme graders is common. Perhaps, however, the clearest evidence of this unreliability was developed in a study by Diederich and two colleagues:

They analyzed the way ten English teachers rated three hundred two-hour compositions by college freshmen in comparison to forty-three other raters: social scientists, natural scientists, writers and editors, lawyers and business executives. The raters were given no standards or criteria for judging the papers, they were merely asked to sort the themes into nine piles in order of general merit, with not less than four percent of the papers in any pile. It was disturbing to find that 94 percent of the papers received seven, eight, or nine, of the possible grades, and that the median correlation between readers was .31. Readers in each field, however, agreed slightly better with the English teachers than with one another. (Braddock, 1963, p. 41)

If there were no more encouraging results than those summarized above, any experiment designed to measure students' skill in written composition would appear to be impractical because of the marker's inability to discriminate consistently. Fortunately, other studies discussing the grading of essays are less pessimistic.

When researchers have taken time to devise ways to mitigate the subjectivity and reduce some of the biases that occur in evaluating essays the unreliability of scores can be decreased appreciably. With markers using a common set of criteria, Buxton reported reader reliability coefficients of .91 and .88, and Kincaid obtained reliabilities ranging from .77 to .91 (Braddock, 1963, p. 42). Finlayson, in his study, "The Reliability of the Marking of Essays," found reliabilities ranging from .79 to .96 in the rating of one-hour papers written by 850 sixth graders in twenty-one Edinburgh primary schools (Braddock, 1963, p. 42). It seems clear, then that in analytic reading, high reader reliabilities are possible when carefully defined criteria are followed.

Similarly, with holistic or impressionistic grading of essays, high inter-rater reliabilities have been obtained when researchers have established standards for the ratings by furnishing readers with copies of sample essays for inspection and discussion and by having readers do some practice marking. Somewhat indicative of the results obtained are the following inter-rater reliabilities: (Black, 1958, p. 179, reliability of .82; Godshalk, Swineford, Coffman, 1966, reliability of .92; and Follam and Anderson, 1967, p. 197, reliability of .95).

Godshalk, Swineford, and Coffman, in obtaining the reader reliability of .92 had students write on five different topics and had each topic read by five different readers. With respect to reliability, Godshalk, et. al. made on the basis of their data and "in the context of previous research" the following generalization:

The reliability of essay scores is primarily a function of the number of different essays and the number of different readings included. If one can include as many as five different topics and have each topic read by as many as five different readers, the reading reliability of the total score may be approximately 0.92 and the score reliability approximately 0.84 for these samples. In contrast, for one topic read by one reader the corresponding figures are 0.40 and 0.25 respectively. The increases which can be achieved by adding topics or readers are dramatically greater than those which can be achieved by lengthening the time per topic or developing special procedures for reading (1966, pp. 39-40).

Procedures to Minimize the Rater Variable. To minimize the effect of the variance among readers the essays were graded by four readers and the total of twelve scores assigned to each student was used as the criterion. All readers contributed to the total score. Furthermore, the investigator, in an effort to establish standards discussed with the raters his concept of writing ability and the writing process. The nature of the essay topics, the conditions of administration, and a general description of the testing sample were also discussed with the raters. Each rater was then furnished with copies of sample essays on all three topics for inspection, discussion, and grading.

II. HOLISTIC VS. ANALYTIC RATING OF ESSAYS

In a comparison study Follman and Anderson (1967) attempted to determine the intra-reliability of each of five different kinds of evaluation procedures and to use those reliability scores as a basis for comparison of the five procedures. Ten themes chosen to

represent the five-point conventional A to F grading continuum were graded by five groups of five raters. "Each rater was randomly assigned to his respective rater group. Each rater group used a different rating procedure. Each rater judged the ten themes independently of the other four raters using the same rating system, as well as independently of the other twenty raters" (p. 196).

Rater Group 1 used the California Essay Scale, Group 2 The Cleveland Composition Rating Scale, Group 3 the Diederich Rating Scale, Group 4 the Follman English Mechanics Guide, and Group 5 the "Everyman's Scale". Of particular interest for this investigation are the correlations of the rating group scores for the different evaluation procedures, especially the correlations of the "Everyman's Scale" with the other four scales:

	Eng. Mech.	Dieder.	Calif.	Cleve.
Every.	.955	.61	.99	.955

Such intercorrelations can be interpreted to support the notion that evaluation systems do in fact measure a substantial number of elements in common.

In a study, "A Comparison of Two Methods of Grading English Compositions," designed to compare the "atomistic" and the "wholistic" methods of grading, Coward (1952) sought answers to two questions "(1) which method of grading is the more reliable, and (2) are the same abilities evaluated by both methods?" It is perhaps sufficient, without summarizing Coward's study, to report her two conclusions:

The wholistic method of reading employed was considerably faster than the atomistic method. It is likely that the reading reliability of the two methods would be about the same if the same amount of time were taken.

There is no evidence in the data of an intrinsic difference in the nature of the abilities evaluated by the two methods of reading employed (p. 93).

That readers, when required to do so, do in fact follow their instructions to "give global judgements" rather than respond to details of an essay, was a conclusion reached by Myers, McConville, and Coffman (1966), and adds further justification to the researcher's use of holistic grading in this investigation. However, with respect to analytical reading Fostvedt (1965) stated that "although teachers of English composition may feel that criteria are important in evaluating themes, there is no evidence of consistency in the employment of such criteria" (p. 111); that is, when grading for coherence and logic, diction, development of ideas, emphasis, and organization through sentence structure and paragraphing. Briefly, Fostvedt first selected criteria for the evaluation of English compositions from nine State and nationwide evaluation scales, and then validated these criteria by finding agreement at the one per cent level in the ranking of the criteria by nine college and twenty-two high school "English experts". When, however, an effort was made to establish reliability of the validated criteria by applying statistical analyses to the numerical ratings of the twenty themes by thirty "English experts" from Montana high schools, Fostvedt found that reliability of the criteria

was not found at the five percent level. Thus, in the studies of Myers, McConville, and Coffman, and of Fostvedt, one notes that in the former readers can successfully judge holistically, while in the latter, "there is no evidence of consistency" in the employment of analytic criteria.

Godshalk, in replying to a criticism of his study, The Measurement of Writing Ability, stated that:

In the matter of analytical versus impressionistic readings, perhaps we did not say enough...the research cited (as well as our own experience in testing programs) indicates that the problem of "halo" reduces the effectiveness of analytical reading... There is, of course, unconscious analysis in all impressionistic readings. For many readers there is some conscious analysis, too, despite instructions to the contrary (1967, p. 87).

Swineford, too, in the same reply stated that "perhaps we should have mentioned that we have worked with analytical ratings and that we have never been able to obtain such ratings sufficiently free from "halo" to perform any useful analysis to the results (1967, p. 87).

Considering the above statements favoring the holistic method of rating compositions, and considering the pressures of extra time and money that analytical reading would require, the researcher adopted holistic rating procedures in this investigation. The method used in rating the compositions is described in the following section.

Description of Holistic Method Used

In using holistic ratings the investigator had first to decide whether to use a three-point, four-point, five-point, seven-point, or ten-point scale, all of which have been used with great frequency by various researchers. Godshalk, Swineford, and Coffman, employing a three-point scale obtained reading reliabilities of .647 for Topic A and .672 for Topic B, but while marking the same topics with a four-point scale obtained reading reliabilities of .714 for Topic A and .685 for Topic B (1966, p. 33). The STEP Essay Tests, however, employ a seven-point scale, whereas in a well known study "A Comparison of Two Methods of Grading English Composition," by Ann Coward (1952) of the Educational Testing Service a ten-point scale was used. Myers, McConville, and Coffman, in a study previously described, employed a four-point scale and obtained with four readers a reliability of .732.

In the matter of score scales, Godshalk of the Educational Testing Service stated "I have used operationally a nine-point scale successfully insofar as observation of process and product can determine success. The readers like it in most cases, too, preferring it to the three-point or four-point scales because they "feel" that it gives greater scope for discriminating judgements and is more "comfortable" as well (1967, p. 87). In the Measurement of Writing Ability, however, Godshalk experimented with a three-point and a four-point scale, and not with a nine-point scale. Godshalk's deciding in favor of the four-point scale was based on the following reasoning: "Very brief

essays do not lend themselves to fine discrimination of quality, in any case. It was our theory that a four-point scale would prevent readers from awarding the "safe" middle score of two on the three-point scale, and so might increase discrimination without creating problems of a trend towards too much analysis" (1967, p. 87). Thus, for purposes of this investigation, a four-point scale was used.

Raters were asked to make global or holistic judgements on each paper, reading rapidly for a total impression, judging "boldly" and decisively. The essay was graded on its impact; that is, holistic grading was intended to be based on subjective, intuitive, overall judgement of a composition. Each rater assigned a rating from 1 to 4 on each essay. The interpretation of these grades was that of Myers, McConville, and Coffman (1966, p. 43), namely:

1. obviously below a reasonable standard,
2. not sufficient promise or competence to be considered in the upper half,
3. clearly competent, promise of effective performance,
4. superior; not perfect but very good; effective.

Readers were requested to judge each paper on its merits without regard to other papers on the same topic, that is, they were instructed not to be concerned with any ideas of a normal distribution of the four scores.

III. PROCEDURE USED FOR GRADING

Though raters did their grading independently (at their homes over a three day period), all essays were rated in a pre-arranged sequence. The three sets of essays were sorted by the investigator using a table of random numbers. A coding sheet on which raters recorded their rating was prepared for each of the three topics. Student papers on STEP 2D were graded first by each rater, the twenty-minute paragraph graded second, and the papers on STEP 2A were graded last. A final tally of the twelve sets of scores (three essays graded by four raters) was made by the investigator and was used as a measure of the student's writing performance, the criterion measure for this investigation.

IV. QUALIFICATIONS OF RATERS

The four raters used in this investigation were all experienced teachers of high school English. Certain defining characteristics of these teachers are shown in Table II.

TABLE II

EXPERIENCE OF RATERS

Raters	Number of Years of Teaching Experience in English			Other Related Experience
	High School	As Head English Teacher	University Teaching	
One	2	1	3	1 year marking English essays through correspondence
Two	8	3		3 years marking grade nine English essays for Dept.ofEd.
Three	4		3	3 years marking grade nine English essays for Dept.ofEd.
Four	6			
(Average)	5			

All readers were deemed competent by the investigator to judge the effectiveness of the students' essays on all three topics. The four raters indicated that they have used a variety of rating procedures with their own students and all have at times used the holistic method of marking. Thus, it seemed reasonable to assume that these four raters, when asked to judge essays holistically, would employ criteria bearing close similarity to the usual goals of composition instruction in secondary schools.

V. SUMMARY OF PROCEDURES FOLLOWED IN ESTABLISHING CRITERION

1. The effects of the assignment variable were minimized by having all students write on each of three essay topics, by requiring written responses consistent with the goals of composition instruction for the tenth-grade, and by using essay topics within students' comprehension and interest.

2. The effects of the writer variable were minimized by requiring students to write essays on three different topics on three different days spaced over a twelve day period, and by using the total score on the three essays as a measure of writing performance.

3. The effects of the rater variable were minimized by having each of the four judges rate each of the three essays, by discussing with the raters the concept of writing ability, the nature of the essay topics, the conditions of administration, the nature of the testing sample, and by furnishing copies of sample essays for inspection, discussion, and grading.

4. Four raters, working independently and using a four-point rating scale, made holistic judgements on each essay. The four raters read all essays in the same prearranged sequence.

5. The sum of the twelve scores (three essays marked by four raters) was used as the criterion measure.

CHAPTER V

DESIGN OF THE STUDY

I. SOURCES OF THE DATA

The study required the administration of five objective measures of writing skills and three essay tests to the sample of six classes of grade ten students. Justification for these particular tests, descriptions of what they are designed to measure, and their reported validities and reliabilities have been presented in Chapter III. Following is a list of the tests used:

(The Standardized Objective Tests Used)

1. Sequential Tests of Educational Progress: Writing, Form 2B
2. Cooperative English Tests: Reading Comprehension, Form 2A
3. Cooperative English Tests: English Expression, Form 2A

(Tests Constructed by the Investigator)

4. Written Expressional Vocabulary
5. Sentence Sensitivity

(Criterion Materials)

6. Essay One: Sequential Tests of Educational Progress: Level 2, Form A
7. Essay Two: Sequential Tests of Educational Progress: Level 2, Form D

II. THE SAMPLE TESTED

The eight measuring instruments (five objective tests, and three free-writing exercises) were administered to six classes of tenth-grade students in one large secondary school in Edmonton, Alberta, a city with a population in excess of 400,000. The school had a total of 917 grade ten students in twenty-nine classes, averaging approximately thirty-one students per class. Complete data were obtained from 151 students - sixty-five boys and eight-six girls.

The six classes were selected so as (1) to allow all students to write the same test on the same day; (2) to provide for a wide range of scores on the criterion measure and on all predictors; and (3) to ensure that the classes were held at different times of the school day so as to permit the investigator to administer all tests.

The school employed homogeneous grouping, based on students' stanine scores on the Grade IX Departmental Examinations, with the average stanine scores for the twenty-nine classes ranging from 3.00 to 8.59. Six classes were selected with the following average stanine scores: 3.80, 5.00, 6.00, 6.94, 7.13, and 8.59.

Other defining properties of the sample are shown in Table III.

TABLE III

RANGES, MEANS, AND STANDARD DEVIATIONS
FOR 151 STUDENTS ON THE CRITERION
VARIABLE AND FIVE PREDICTOR VARIABLES

Variable	Maximum Range	Obtained Range	Mean	Standard Deviation
Written Composition	12 - 48	14 - 44	28.83	6.37
Written Exp- ressional Vocabulary	0 - 94	27 - 83	53.72	11.64
Sentence Sensitivity	0 - 79	25 - 64	48.71	7.90
English Effective- ness	0 - 90	33 - 76	56.17	9.34
STEP: Writing	0 - 60	25 - 55	40.19	6.37
Reading Com- prehension	0 - 120	35 - 117	84.87	16.69
Intelligence		96 - 150	125.52	12.77

III. TEST ADMINISTRATION

The eight tests were administered by the investigator to each of the six grade ten classes during regular forty-three-minute English class periods. A total of nine days (54 class periods) during the months of February and March, 1969, was used for this testing. The tests were administered in the following sequence:

Test (2)	40 minutes
Test (6)	35 minutes
Test (1, Part I)	35 minutes
Test (2, Part II)	35 minutes
Test (8)	20 minutes
Test (3)	40 minutes
Test (7)	35 minutes
Test (4)	30 minutes
Test (5)	30 minutes

IV. TREATMENT OF THE DATA

Standard statistical techniques were used for the analysis of data.

The item analysis carried out for the Pilot Study has already been discussed on pages 87-90.

Using the facilities of the Division of Educational Research Services of the University of Alberta, a three way analysis of variance with repeated measures on the last two factors was performed for the purpose of estimating the reliability of reading and the reliability of the total essay score. A series of correlations between essay scores of different raters for different topics was examined to illustrate the influence of the number of readings upon the reliability of the criterion measure.

Pearson product-moment correlations, multiple correlations, and multiple linear regression analyses were used to test the main hypotheses on the relations between the criterion and the five objective tests.

As its name implies, multiple linear regression seeks only to clarify whether or not a critical variable, when added to a linear expression, significantly reduces the criterion error sum of squares. The general approach is to express a vector of criterion variable data as a linear combination of a set of predictor vectors:

$$\hat{Y} = A_1X_1 + A_2X_2 + \dots \dots \dots A_nX_n + E$$

where:

Y is a vector of criterion variable data ($N \times 1$)

$X(i)$ are vectors of predictor variable data ($N \times 1$)

$A(i)$ are unknown weights associated with the predictor vectors

E is an error or residual vector

N is the number of observations

The problem is to find a set of weights ("least square weights") which minimize the error sum of squares (ESS) between the predicted criterion, \hat{Y} , and the measured values of Y . The error sum of squares is calculated over the N individuals for whom scores are available, as follows:

$$ESS = \sum_{i=1}^N (Y_i - Y_{1i})^2$$

The observed product moment correlation, R_1 , between Y and \hat{Y}_1 is a measure of the goodness of fit between observed and predicted values of the criterion. Its square, called the squared multiple correlation (RSQ), represents the proportion of the variance of criterion accounted for by the linear equation predicting \hat{Y}_1 , which is called, say, Model 1.

To investigate, in the presence of other variables, the effect of one particular variable, a new model, Model 2, is written. For example to test the effect of variable three $X(3)$, Model 2 is written such that:

$$\hat{Y}_2 = A_1X_1 + A_2X_2 + A_4X_4 + \dots + A_nX_n + E_2$$

This equation, since it restricts variable three, is called the restricted model, whereas, the equation using all the possible predictors for the study is referred to as the unrestricted (general) model. From the restricted model a squared multiple correlation (R_2^2), which will be less than or equal to R_1^2 , is calculated.

Thus the significance of the contribution of any one variable $X(i)$ in the presence of others can be tested by calculating the F ratio as follows:

$$F = \frac{(R_1^2 - R_2^2) / df_1}{(1 - R_1^2) / df_2}$$

where:

R_1^2 = the squared multiple correlation from the full model

R_2^2 = the squared multiple correlation from the restricted model

$df_1 = (m_1 - m_2)$ with m_1 being the number of unknown weights associated with the full model, and m_2 the number of unknown weights associated with the restricted model

$df_2 = (N - m_2)$, with N being the number of observations

In a similar manner, all n predictor variables of an experiment can be investigated (Bottenberg and Ward, 1963).

CHAPTER VI

RESULTS: THE CRITERION MEASURE

Basic to this investigation was the importance of having a highly reliable score of writing performance to serve as a sound criterion for validating the five objective tests. The methods used and the efforts made to establish such a criterion were discussed in Chapter IV. Briefly, it was suggested that there is a writer variable, an assignment variable, and a rater variable and that the unreliability that may result from these variables can be minimized by using the total score (three topics each marked by four raters) as a measure of writing performance. In other words, it was suggested that reliability may be a function of the number of essays written by the students and the number of raters grading these essays. It was also suggested that high reading reliability may be obtained when using a holistic method of marking. In this study each of four raters assigned grades (from 1 to 4) on each of three essays written by 151 students. The means, standard deviations, intercorrelations of the sets of scores, summaries, and statistical treatment of the data are presented in this chapter.

Table IV shows the means and standard deviations of scores of the four readings on each of the three essay topics. The mean essay score for the 151 students was 28.80, with the maximum range

TABLE IV
 MEANS AND STANDARD DEVIATIONS OF SCORES OF THE
 FOUR READINGS ON EACH OF THREE ESSAY TOPICS
 (N = 151)

Essays		Raters				Total Score
		One	Two	Three	Four	
Essay One (STEP 2A)	Mean	1.99	2.48	2.41	2.84	9.72
	S.D.	.83	.79	.69	.84	2.75
Essay Two (STEP 2D)	Mean	1.92	2.33	2.55	2.73	9.52
	S.D.	.83	.93	.68	.86	2.75
Essay Three (Paragraph)	Mean	2.21	2.28	2.42	2.66	9.57
	S.D.	.65	.79	.68	.83	2.24
Total	Mean	6.10	7.07	7.37	8.22	28.80
	S.D.	1.87	1.76	1.63	2.06	6.40

of scores being from 12 to 48. The total mean score of assigned grades by raters ranged from 6.10 by rater one to 8.22 by rater two. The total mean scores assigned to individual topics ranged from 9.52 for Essay Two to 9.72 for Essay One. The mean scores for each of the three essays and for each of the four raters are also given in Table IV.

The increase in reliability due to the number of topics and the number of raters is demonstrated through the sequence of Tables V to VIII. Table V shows the correlations between the readings of each of four raters on each of the three essay topics. It is shown, for example, that on Essay One the correlations range from .29 between raters one and two to .55 between raters one and four. The average of the six correlations on Essay One is .46, which may be interpreted as being approximately equal to the reliability of a single rating. Using this reliability of a single rating, (.46), in the following formula one can estimate the reliability of the mean of four ratings, which for the present example is:

$$r_k = \frac{kr_1}{1 + (k-1)r_1}$$

$$= \frac{4(.46)}{1 + (4-1)(.46)} = .77$$

This means that if a second group of four raters as competent as the first group were to read the same essay, it might be expected that their scores would produce a correlation of approximately .77.

TABLE V

CORRELATIONS BETWEEN THE SINGLE READINGS OF EACH
OF FOUR READERS ON EACH OF THREE ESSAY TOPICS

	Raters	One	Two	Three	Four
Essay One (STEP 2A)	One		.29	.48	.55
	Two	.29		.43	.50
	Three	.49	.43		.48
	Four	.55	.50	.48	
	Total	.75*	.72*	.75*	.83*
Essay Two (STEP 2D)	One		.54	.61	.66
	Two	.54		.54	.61
	Three	.61	.54		.61
	Four	.66	.61	.61	
	Total	.83*	.82*	.80*	.86*
Essay Three (Paragraph)	One		.30	.48	.48
	Two	.30		.34	.51
	Three	.48	.34		.47
	Four	.48	.51	.47	
	Total	.71*	.73*	.73*	.83*

*Spuriously high because part is included in the total.

Thus, for one rater on Essay One the reading reliability is .46, but for four raters on the same essay, the reading reliability is .77. Similar reliabilities for Essays Two and Three are shown in Table VIII.

Table VI shows the correlations between the readings of each of four raters on each of the three pairs of essays, namely Essays One and Two, One and Three, and Two and Three. The correlations of these summed scores for each of the four raters on each of the three combinations of essays are shown in Table VI. For example, the correlations on the combination of Essays One and Three range from .39 between raters one and two to .62 between raters two and four. The mean correlation among raters on this combination of Essays One and Three is .53. This average correlation may be interpreted as the reliability of one rater on two topics. From this reliability of a single rating the reliability of the mean of four ratings on two essays may be obtained. The result is a correlation of .81. This means that if a second group of four raters were chosen to mark the two essays (Essay One and Essay Three) and the two scores were added, the correlation between the scores for the second group of raters and the first group would be approximately .81. Thus, the reliability of reading for one rater on Essays One and Three is .53, but for four raters on the two essays the reliability is .81. Similar reliabilities for the other combinations of two essays are shown in Table VIII.

TABLE VI

CORRELATIONS BETWEEN READINGS OF EACH OF FOUR READERS
ON EACH OF THREE PAIRS OF ESSAYS

	Raters	One	Two	Three	Four
(Essays One & Two)	One		.58	.62	.70
	Two	.58		.68	.71
	Three	.62	.68		.67
	Four	.70	.71	.67	
	Total	.84*	.85*	.84*	.90*
(Essays One & Three)	One		.39	.54	.61
	Two	.39		.47	.62
	Three	.54	.47		.57
	Four	.61	.62	.57	
	Total	.78*	.77*	.78*	.88*
(Essays Two & Three)	One		.60	.66	.72
	Two	.60		.59	.69
	Three	.66	.59		.60
	Four	.72	.69	.60	
	Total	.86*	.84*	.81*	.88*

*Spuriously high because part is included in the total.

Table VII shows the correlations between readings of each of four readers on three essays. Here, the range of correlations is from .62 for raters one and two to .75 for raters two and four. The average of the six correlations is .68 and may be interpreted as the reliability of one rater on three essays. From this reliability of .68 the reliability of four raters on three topics is estimated to be .89. This means that if a second group of four raters were chosen and the essays were read again, it might be expected that the two sets of scores would produce a correlation of approximately .89.

In summary, Tables V, VI, and VII show the correlations between the readings of four readers on each of three essays, on pairs of essays, and finally on the total of three essays. As presented above, these correlations were used to indicate the reliability of a single rater and that of four raters, on one, two, and three topics. Table VIII summarizes these reliabilities and demonstrates the fact that the reading reliability of the criterion increases with both the number of essays and the number of raters, and for the total criterion score used in this study is .89.

In addition to demonstrating this fact, the essay scores as assigned by individual raters and for individual topics may be used to illustrate other points of relevance for the assessment of writing ability. To obtain a more comprehensive picture of the influence of such factors as the rater and topic upon total essay score, as well as to provide an estimate of reading and total score reliability, an analysis of variance was performed on the ratings.

TABLE VII
CORRELATIONS BETWEEN READINGS OF EACH OF FOUR RATERS
ON THREE ESSAY TOPICS

	Raters	One	Two	Three	Four
Total Essay Score	One		.62	.65	.72
	Two	.62		.67	.75
	Three	.65	.67		.66
	Four	.72	.75	.66	
	Total	.85*	.86*	.84*	.91*

*Spuriously high because part is included in the total.

TABLE VIII

TABLE SHOWING INCREASES IN CORRELATIONS WITH THE ADDITION
OF ESSAYS AND RATERS

Reliability of one rater
on one essay

Essay One	.46
Essay Two	.60
Essay Three	.43

Reliability of one rater
on two essays

Essays One & Two	.66
Essays One & Three	.64
Essays Two & Three	.53

Reliability of one rater
on three essays

Essays One, Two, and Three	.68
-------------------------------	-----

Estimate of reliability of
four raters on one essay

Essay One	.77
Essay Two	.86
Essay Three	.75

Estimate of reliability of
four raters on two essays

Essays One & Two	.88
Essays One & Three	.88
Essays Two & Three	.81

Estimate of reliability of
four raters on three essays

Essays, One, Two, and Three	.89
--------------------------------	-----

A three-way analysis of variance program was used (Winer, 1962, pp. 312-337), with the main interest being in Factor B (topics) and in Factor C (raters). There were repeated observations on these two factors. Factor A was assumed to have only one level and thus accounted for no variance in this design. Table IX summarizes the results of this analysis of variance.

The reading reliability was estimated by dividing the mean square for students minus the mean square for error by the mean square for students. From Table IX the mean square for students is given by $SS_{\text{between subjects}}/df$. The mean square for error was estimated by the formula (Winer, 1962, p. 322):

$$MS_{\text{error}} = \frac{SS_b \times \text{subj.w.groups} + SS_c \times \text{subj.w.groups} + SS_{bc} \times \text{subj.w.groups}}{p(n-1)(qr-1)}$$

$$= \frac{184.56 + 157.27 + 261.71}{(150)(11)} = 0.37$$

Specifically,

$$\text{reading reliability} = \frac{\text{students mean square} - \text{error mean square}}{\text{students mean square}}$$

$$= \frac{3.40 - 0.37}{3.40} = 0.89$$

The reading reliability means that if a second group of four readers as competent as the first group were to read the same essays, it might be expected that the two sets of total scores would produce a correlation of .89.

TABLE IX
SUMMARY OF ANALYSIS OF VARIANCE OF ESSAY READING

Source of Variation	Sum Squares	Degrees of Freedom	Mean Square	F	P
Between Subjects	510.18	150	3.40		
A	0.00	0	0.00	0.0	
Subj. w. groups	510.18	150	3.40		
Within Subjects	733.33	1661			
B (topics)	0.77	2	0.38	0.63	0.53
AB	0.00	0	0.00		
B x subj. w. groups	184.56	300	0.61		
C (raters)	114.73	3	38.24	109.43	0.0
AC	0.00	0	0.00	0.0	
C x subj. w. groups	157.26	450	0.34		
BC (topics x raters)	14.28	6	2.38	8.19	0.0
ABC	0.00	0	0.00	0.00	
BC x subj. w. groups	261.71	900	0.29		

The score reliability of the total essay score was estimated by using the mean square for the interactions of students and topics as the error term - in terms of Table IX, this is the mean square for B x subj. w. groups. Thus,

$$\text{score reliability} = \frac{\text{students mean square} - \text{student} \times \text{topics mean square}}{\text{students mean square}}$$

$$= \frac{3.40 - 0.615}{3.40} = 0.82$$

The score reliability is an estimate of the correlation to be expected if the students were to write three more essays on three new topics and if the essays were read by four new readers.

Turning to the main effects of Factors B and C, Table IX indicates a significant rater effect. This may be taken to mean that scores were higher or lower for some of the four raters than they were for others. This finding is indicative of the fact that a student's essay score may be a function not only of writing skill but also of the particular rater. The variation in raters could be controlled either through totalling across raters or by the transformation of scores to make them comparable.

The main effects of Factor B (topics) was not significant. However, in view of the significant BC interaction (topics x raters) it is more meaningful to consider differences between means of scores across topics for each rater separately. The means and standard deviations of scores for the four readings of each of three essay topics are given in Table IV. A geometric representation of

these means is given in Figure I, which may be viewed in interpreting the rater x topic interaction. Inspection of Figure I indicates that while raters two and four on the average treated the topics in essentially the same way, other raters differed on the average in their assignment of scores by topic; for example rater one assigned higher scores to topic three and lower scores to topic two, whereas the reverse was true for rater three. Since the topic does seem to influence the scores assigned by certain raters, care should be taken in the assessment of writing performance to counteract bias introduced by the topics presented. The method used for the provision of the present criterion could be used, where the score is the sum across a number of topics.

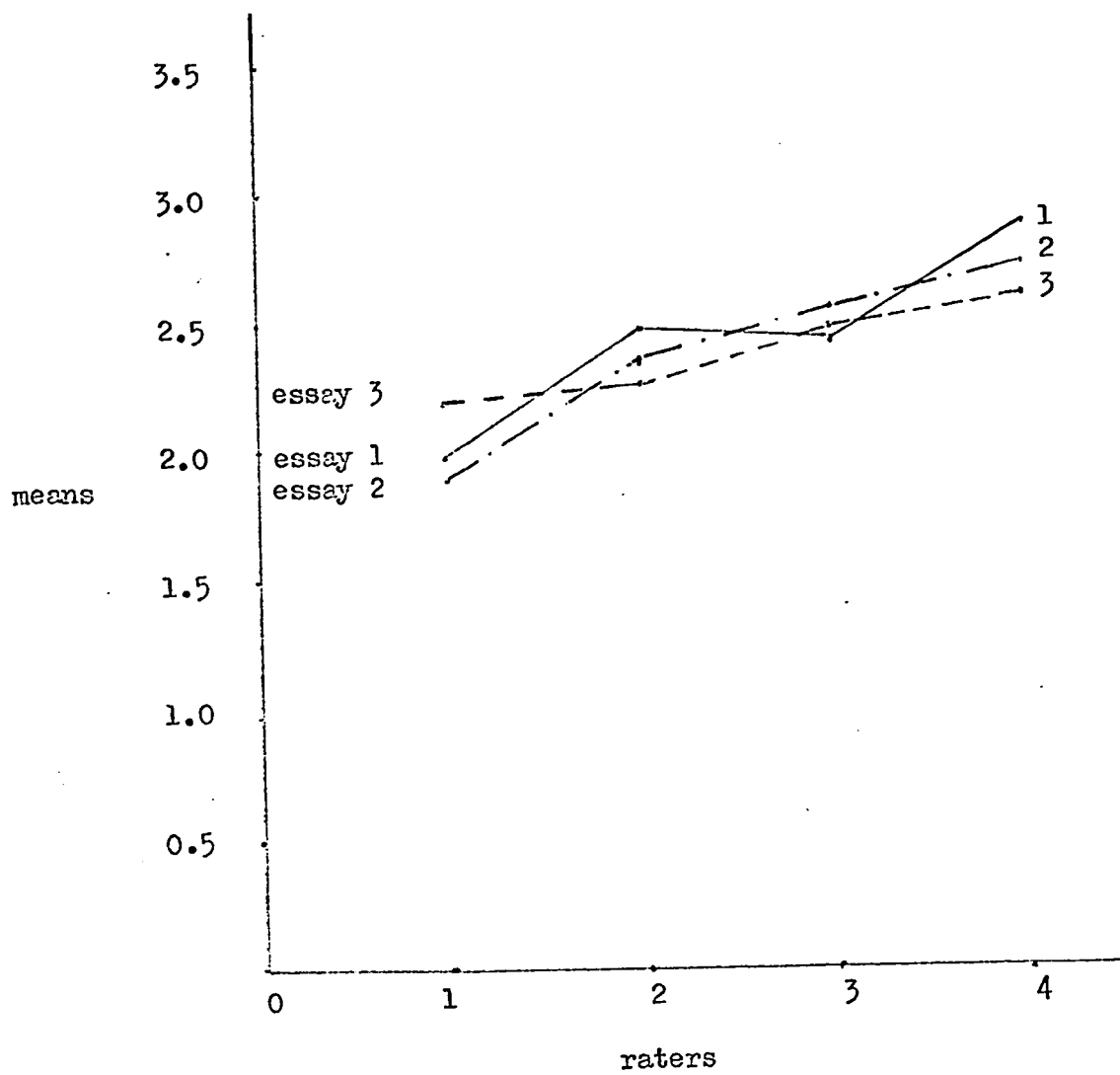


Figure 1: Comparison of Means Showing the Rater by Essay Interaction

CHAPTER VII

RESULTS: RELATIONS BETWEEN PREDICTORS AND THE CRITERION

With the criterion having been defined, this chapter will present the results on the relations between the objective tests and the criterion. Hypotheses 2, 3, 4, 5, and 6 guide this presentation of results.

Hypothesis 2

There will be a significant positive correlation between the total measure of writing performance and the following predictor variables (taken individually and in combinations of two and three):

Written Expressional Vocabulary
Sentence Sensitivity
English Expression
Effectiveness
Mechanics
STEP: Writing
Reading Comprehension
Vocabulary
Comprehension

Table X shows the Pearson product-moment correlations of the five objective tests with the total essay score and with the total scores on each of the three essays. As hypothesized, all correlations with the total essay score are significant. With a sample size of 151, a chosen alpha of 0.05, and 149 degrees of freedom, a correlation of 0.15 was needed for significance. The validity coefficients of the five objective tests with the total essay score are .589, .618, .622, .569, and .588*. Though a small range (.569 to

*Correlation coefficients are here used as indices of validity

TABLE X

CORRELATIONS OF OBJECTIVE TESTS WITH THE SUM OF FOUR
SCORES ON EACH ESSAY AND WITH THE TOTAL ESSAY SCORE

Objective Tests	Essay One STEP 2A	Essay Two STEP 2D	Essay Three Paragraph	Total Essay Score
Written Expressional Vocabulary	.532	.453	.553	.589
Sentence Sensitivity	.532	.540	.518	.618
English Expression	.580	.504	.516	.622
Effectiveness	.301	.316	.345	.375
Mechanics	.598	.475	.491	.606
STEP: Writing	.471	.507	.478	.569
Part I	.288	.355	.327	.380
Part II	.529	.530	.505	.610
Reading Comprehension	.474	.489	.562	.588
Vocabulary	.464	.478	.539	.572
Comprehension	.422	.445	.507	.531

.622) exists between these five coefficients, with the least valid being the STEP: Writing, the two tests constructed by the investigator rank second and third (.618 for the Sentence Sensitivity test and .589 for the Written Expressional Vocabulary test). In view of these coefficients it appears that all five tests are valid, and give support to the investigator's conceptualization of some of the components of writing ability as outlined in Chapters II and III.

In addition to obtaining the validity coefficients of the individual objective tests, it was the purpose of this investigation to determine the relative validities of selected combinations of the objective tests as predictors of an essay measure of writing performance (the total score of the three essays). With the five objective tests used in this investigation, there are ten possible combinations when the tests are used in pairs. The validity coefficients for these combinations are presented in Table XI. These coefficients are multiple correlation indices and thus represent the correlation between the criterion variable and a weighted sum of the two predictor variables.

For ease and clarity of presentation all variables, both criterion and predictor, were assigned the following test numbers:

1. Essay Three: Paragraph
2. Total Essay Score
3. Total Score of Essay One: STEP 2A and Essay Two: STEP 2D

TABLE XI
MULTIPLE CORRELATIONS FOR COMBINATIONS OF
PREDICTORS WITH TOTAL ESSAY SCORE AS CRITERION

Test Combinations (in two's)	Multiple Correlation
5, (6, 7)	.702
5, (9, 10)	.702
4, 5	.700
(6, 7), (9, 10)	.689
4, (6, 7)	.689
(6, 7), 8	.677
5, 8	.677
4, 8	.670
8, (9, 10)	.657
4, (9, 10)	.639
Test Combinations (in three's)	Multiple Correlation
5, (6, 7), (9, 10)	.741
4, 5, (6, 7)	.737
5, (6, 7), 8	.722
4, 5, 8	.721
4, (6, 7), 8	.718
4, 5, (9, 10)	.717
5, 8, (9, 10)	.716
(6, 7), 8, (9, 10)	.713
4, 8, (9, 10)	.684
Test Combinations (in four's)	Multiple Correlation
4, 5, (6, 7), (9, 10)	.751
5, (6, 7), 8, (9, 10)	.749
4, 5, (6, 7), 8	.749
4, 5, 8, (9, 10)	.730
4, (6, 7), 8, (9, 10)	.728
Test Combination (all five tests)	Multiple Correlation
4, 5, (6, 7), 8, (9, 10)	.758

4. Written Expressional Vocabulary
5. Sentence Sensitivity
6. Effectiveness (part of English Expression)
7. Mechanics (part of English Expression)
8. STEP: Writing
9. Vocabulary (part of Reading Comprehension)
10. Comprehension (part of Reading Comprehension)
11. Sex
12. Intelligence

Thus, the expression, objective tests 4, 5, means the combination of the Written Expressional Vocabulary Test and the Sentence Sensitivity Test.

In contrast to the range of validity coefficients, .569 for test 8 to .622 for test (6, 7), for the individual objective tests, the range for the two-test combinations is from .639 for tests 4, (9, 10) to .702 for tests 5, (6, 7). The least valid combination (.639) is that of the Written Expressional Vocabulary Test and the Reading Comprehension Test. However, these two tests would probably not be used in a combination since the Reading Comprehension Test includes a vocabulary test which correlates .69 with the Written Expressional Vocabulary Test. The high correlation of .72 between the Written Expressional Vocabulary Test and the Reading Comprehension Test accounts to some extent for the relatively low correlation of the combination with the total essay score.

The three sets of two-test combinations with the highest correlations (.700 to .702) with the criterion include test 5, the investigator's Sentence Sensitivity Test. This finding suggests that the characteristics of sentences described in Chapters II and III as being theoretically central to written composition are indeed valid components of writing ability. Further support for the investigator's concept of writing ability is that the combination of the two tests constructed by the investigator, Written Expressional Vocabulary and Sentence Sensitivity, has a multiple correlation of .700 with the three essay criterion.

With the three-test combinations the range of multiple correlations is from .684 for tests 4, 8, (9, 10) to .741 for tests 5, (6, 7) (9, 10). Again, the Sentence Sensitivity Test is in the four combinations with the highest validity coefficients. Finally, for four-test combinations the range of multiple correlations, as presented in Table XI is from .728 to .751, while the multiple correlation for all five tests with the total essay score is .758.

Though all the above correlations are relatively high and are encouraging, it is more important for this investigation to examine the data in terms of Hypothesis 3 on the unique contribution of the Sentence Sensitivity Test and the Written Expressional Vocabulary Test to the total essay score variance.

Hypothesis 3

- (a) The combination of all five objective tests will account for the highest proportion of total essay score variance, relative to other examined combinations.
- (b) The unique contribution of students' scores on the Sentence Sensitivity Test to this total essay score variance will be significant.
- (c) The unique contribution of students' scores on the Written Expressional Vocabulary Test to this total essay score variance will be significant.

Multiple linear regression analysis was used to test this hypothesis. Table XII shows the correlation matrix for the five predictor variables and the criterion variable, total essay score. These correlations were used to determine the "least square weights" associated with the unrestricted and restricted models, as discussed in Chapter V.

The regression equation for the unrestricted model, the proportion of essay score variance for which the model accounts, and the significance of the contribution of the different predictor variables to this variance are shown in Table XIII. As hypothesized, the combination of all five objective tests accounts for the highest proportion of total essay score variance, when compared with other combinations tested. As shown in Table XI, the highest proportion of total essay score variance accounted for by combinations of two,

TABLE XII

INTERCORRELATIONS AMONG PREDICTOR
VARIABLES AND THE CRITERION VARIABLE

(N = 151)

[illegible]

TABLE XIII
CONTRIBUTION OF PREDICTOR VARIABLES WITH
TOTAL ESSAY SCORE AS CRITERION

	Restriction	R_u^2	R_r^2	df	F-Ratio	Probab- ility*
X(4)	Written Exp. Voc.	.575	.561	1/143	4.45	0.034
X(5)	Sentence Sensitivity	.575	.529	1/143	15.50	<0.001
X(6)	Effectiveness	.575	.565	1/143	3.49	0.063
X(7)	Mechanics	.575	.540	1/143	11.79	<0.001
X(6) + (7)	English Expression	.575	.532	2/143	7.24	0.001
X(8)	STEP: Writing	.575	.564	1/143	3.84	0.051
X(9)	Vocabulary	.575	.563	1/143	4.05	0.046
X(10)	Comprehension	.575	.575	1/143	0.02	0.878
X(9)+X(10)	Reading Comprehension	.575	.561	2/143	2.39	0.094

REGRESSION EQUATION FOR UNRESTRICTED MODEL

$$Y = 0.096X(4) + 0.236X(5) - 0.245X(6) + 0.215X(7) + 0.149X(8) \\ + 0.150X(9) - 0.007X(10) + 2.447$$

*Level of significance = 0.05.

three, or four is .564, the proportion accounted for by the first listed four-test combination (Table XI, $R = .751$). This proportion is less than that which accounted for by the five-test combination, the latter being .575 as indicated by the R_u^2 of Table XIII. To examine the unique contribution of the Sentence Sensitivity test to the total essay score variance, the regression equation for the restricted model was:

$$\hat{Y} = 0.121X(4) - 0.158X(6) + 0.274X(7) + 0.230X(8) + 0.138X(9) \\ + 0.007X(10) + 0.431$$

As postulated by Hypothesis 3 (b), students' scores on the Sentence Sensitivity Test contribute significantly to the variance of total essay scores. The F-ratio, 15.50, is highly significant ($p < .001$), which means that taking into account the correlations of Written Expressional Vocabulary, English Expression (effectiveness, and mechanics), STEP: Writing, and Reading Comprehension (vocabulary, and comprehension) with the criterion variable and the inter-correlations of the predictor variables with one another, Sentence Sensitivity contributes to the regression equation predicting students' composition scores. In other words, knowledge of students' scores on the Sentence Sensitivity Test enables one to make a significantly better prediction of students' composition scores than one could make by just knowing students' scores on the four objective tests of Written Expressional Vocabulary, English Expression, STEP: Writing, and Reading Comprehension.

The regression equation to test Hypothesis 3 (c) or the unique contribution of the Written Expressional Vocabulary Test was:

$$\hat{Y} = 0.251X(5) - 0.227X(6) + 0.237X(7) + 0.149X(8) + 0.206X(9) + 0.017X(10) - 2.501$$

As postulated by this hypothesis, students' scores on the Written Expressional Vocabulary Test contribute significantly to the variance of total essay scores. The F-ratio for the unrestricted and restricted models is 4.45, with a probability of 0.034. Thus, students' Written Expressional Vocabulary scores contribute significantly to the variance of total essay scores when cognizance has been taken of the effects of scores on the other four objective tests. On the basis of this analysis the hypothesis was accepted.

Table XIII shows that four other variables, Mechanics ($p = .001$), English Expression ($p = .001$), STEP: Writing ($p = .051$), and the Vocabulary section of the Reading Comprehension Test ($p = .046$) contributed significantly to the variance of total essay scores. It should be remembered, however, that the results of these tests for a particular variable have meaning only in the presence of the variables included in the regression equation.

Hypothesis 4

- (a) There will be a significant positive correlation between scores on a two-essay criterion and the following predictor variables (taken individually and in combinations of two and three):

Written Expressional Vocabulary
 Sentence Sensitivity
 English Expression
 Effectiveness
 Mechanics
 STEP: Writing
 Reading Comprehension
 Vocabulary
 Comprehension
 A Twenty-minute Paragraph

- (b) The addition of a twenty-minute paragraph will increase the multiple correlation between scores on the two-essay criterion and each of the two- and three-objective test combinations.

This aspect of the investigation was an attempt to determine the effects of the addition of a twenty-minute paragraph to the selected combinations of predictors. To do so, first necessitated changing the three essay criterion to a two-essay criterion and using the paragraph as an extra predictor. This two-essay criterion is now less reliable than the total essay score on three essays (see correlations, Table VIII). Table XIV shows the correlations of all predictors with the new two-essay criterion. The correlations range from .517 for the Reading Comprehension Test to .586 for the Sentence Sensitivity Test, with the correlation for the paragraph with the criterion being .617.

Table XV shows the correlations for the ten possible combinations of the five objective tests taken in pairs. The correlations range from .569 (for the Written Expressional Vocabulary Test plus

TABLE XIV

CORRELATIONS OF SIX PREDICTORS WITH TOTAL SCORE
ON A TWO ESSAY CRITERION (ESSAY ONE: STEP 2A;
AND ESSAY TWO; STEP 2D)

Predictors	Total Score on Two Essays
Paragraph	.617
Written Expressional Voc.	.530
Sentence Sensitivity	.586
English Expression	.579
Effectiveness	.330
Mechanics	.576
STEP: Writing	.534
Reading Comprehension	.517
Vocabulary	.513
Comprehension	.467

TABLE XV

MULTIPLE CORRELATIONS FOR SELECTED COMBINATIONS OF
PREDICTORS WITH A TWO-ESSAY CRITERION (ESSAY ONE:
STEP 2A; AND ESSAY TWO: STEP 2D)

Test Combinations (in twos)		Addition of Paragraph to Same Test Combinations	
Tests	Validity	Tests	Validity
5,(6,7)	.662	1,5,(6,7)	.721
5,(9,10)	.645	1,5,(9,10)	.704
5,4	.645	1,5,4	.704
8,(6,7)	.636	1,8,(6,7)	.711
4,(6,7)	.635	1,4,(6,7)	.703
5,8	.635	1,5,8	.708
(6,7),(9,10)	.634	1,(6,7),(9,10)	.702
4,8	.612	1,4,8	.689
8,(9,10)	.597	1,8,(9,10)	.684
4,(9,10)	.569	1,4,(9,10)	.665

Test Combinations (in threes)		Addition of Paragraph to Same Test Combinations	
Tests	Validity	Tests	Validity
4,5,(6,7)	.687	1,4,5,(6,7)	.728
5,(6,7),8	.681	1,5,(6,7),8	.731
5,(6,7),(9,10)	.677	1,5,(6,7),(9,10)	.731
4,5,8	.666	1,4,5,8	.715
4,(6,7),8	.666	1,4,(6,7),8	.718
5,8,(9,10)	.661	1,5,8,(9,10)	.713
(6,7),8,(9,10)	.661	1,(6,7),8,(9,10)	.716
4,5,(9,10)	.658	1,4,5,(9,10)	.707
4,8,(9,10)	.621	1,4,8,(9,10)	.690

the Reading Comprehension Test) to .662 (for the Sentence Sensitivity Test plus the English Expression Test). As shown on the right hand side of Table XV the range of these correlations changes from (.569 to .662) to (.665 to .721) with the addition of the paragraph to the predictor variables. Although the combination of the Sentence Sensitivity Test and the English Expression Test is the most valid (.662) two-set combination of predictors, the correlation of this combination with the criterion increases from .662 to .721 when the twenty-minute paragraph is added as a predictor. In every instance, the paragraph adds to the prediction.

The comparisons for the combinations of predictors taken three at a time, first without the paragraph and then with the paragraph, are also shown in Table XV. The most valid (.687) three-test combination of predictors includes tests 4, 5, and (6,7). Tests 4 and 5 are the tests constructed by the investigator. With the addition of the paragraph this validity coefficient increases from .687 to .728. Again, as with the two-test combinations, the addition of the paragraph increases the prediction of every three-test combination. Finally, as a matter of interest, the multiple correlation for all five objective tests with the two-essay criterion is increased to .739 from .707 with the addition of the twenty-minute paragraph.

Hypothesis 5

- (a) The combination of all five objective tests plus the paragraph will account for the highest proportion of variance on the two essay criterion, relative to other combinations.
- (b) The unique contribution of students' scores on the paragraph to this two-essay score variance will be significant.

Hypothesis 5 was also examined through the use of multiple regression analysis. The proportion of the two-essay variance accounted for by the five objective tests plus the paragraph is indicated in Table XVI as R_u^2 . This proportion of .546 is higher than that which is accounted for by either of the two- or three-test combinations shown in Table XV. The highest proportion of the two-essay variance accounted for by such combinations is .534, as accounted for by the second- and third-listed three-test combination (Table XV, $R = .731$).

Table XVI also presents the R_r^2 for the restricted models used to examine the unique contributions of each of the five objective tests and the paragraph to the predicted two-essay score variance. The regression equation used to test the contribution of the paragraph was:

$$\hat{Y} = 0.059X(4) + 0.178X(5) - 0.201X(6) + 0.170X(7) \\ + 0.116X(8) + 0.110X(9) - 0.029X(10) - 2.532$$

The F-Ratio of 14.48 and the probability of <.001 are shown in Table XVI.

TABLE XVI

CONTRIBUTION OF PREDICTOR VARIABLES WITH TOTAL SCORE ON TWO
ESSAYS AS CRITERION AND WITH PARAGRAPH USED AS A PREDICTOR

	Restriction	R_u^2	R_r^2	df	F-Ratio	Probab- ility*
X(1)	Paragraph	.546	.499	1/142	14.48	<0.001
X(4)	Written Exp. Vocabulary	.546	.542	1/142	1.03	0.310
X(5)	Sentence Sensitivity	.546	.518	1/142	8.65	0.003
X(6)	Effectiveness	.546	.538	1/142	2.51	0.114
X(7)	Mechanics	.546	.519	1/142	8.35	0.004
X(6)+X(7)	English Expression	.546	.513	2/142	5.10	0.007
X(8)	STEP: Writing	.546	.537	1/142	2.64	0.106
X(9)	Vocabulary	.546	.541	1/142	1.48	0.225
X(10)	Comprehension	.546	.541	1/142	1.33	0.250
X(9)+X(10)	Reading Comprehension	.546	.541	2/142	0.76	0.466

REGRESSION EQUATION FOR UNRESTRICTED MODEL

$$\begin{aligned}
 Y = & 0.612X(1) + 0.029X(4) + 0.138X(5) - 0.164X(6) \\
 & + 0.145X(7) + 0.094X(8) + 0.069X(9) - 0.032X(10) \\
 & - 2.549
 \end{aligned}$$

*Level of significance = 0.05.

As was postulated by Hypothesis 5 (b), students' scores on the paragraph contribute significantly to the variance of total scores on the two-essay criterion. This means that taking into account the correlations of tests 1, 4, 5, (6, 7), 8, (9, 10) with the criterion variable and the intercorrelations of the predictor variables with one another, the addition of the paragraph contributes significantly to the prediction of students' composition scores.

Four of the ten variables (paragraph, sentence sensitivity, effectiveness, and English expression) in Table XVI are shown to make significant contributions to the prediction of scores on the two-essay criterion. The Sentence Sensitivity Test has an F-ratio of 8.65 and a probability of 0.003, indicating that in the presence of the nine other variables (including the paragraph) the test contributes significantly to the prediction of composition scores. The effectiveness part of the English Expression Test and the English Expression Test itself also make significant contributions. Variables with a probability greater than 0.05 do not in themselves contribute significantly to the prediction of essay scores when used in the specific combinations outlined by Table XVI.

Hypothesis 6

- (a) The unique contribution of students' scores on the Sentence Sensitivity Test to the total essay score variance (as in Hypothesis 3 b) will remain significant when sex and intelligence are included as predictor variables.

- (b) The unique contribution of students' scores on the Written Expressional Vocabulary Test to the total essay score variance (as in Hypothesis 3 c) will remain significant when sex and intelligence are included as predictor variables.
- (c) The unique contribution of students' scores on the paragraph to the two essay score variance (as in Hypothesis 5 b) will remain significant when sex and intelligence are included as predictor variables.

A final purpose of this investigation was to determine whether the tests of Sentence Sensitivity and Written Expressional Vocabulary, and whether the twenty-minute paragraph would make significant contributions to the variance of essay scores when knowledge of students' sex and intelligence were inserted into the prediction equations. Using the total essay score on the three essays as a criterion, Table XVII shows that the total variance of composition scores accounted for by the combination of all predictors was 57.6%. As shown in Table XIII, without the use of knowledge of sex and intelligence the total variance accounted for by all predictors was 57.5%. The increase in prediction is thus negligible. Sex and intelligence themselves, have F-ratios of 0.12 and 0.06, and probabilities of .721 and .796 respectively. Thus, in the presence of the variables considered in this study, sex and intelligence do not make significant contributions to the prediction of composition scores.

TABLE XVII

CONTRIBUTION OF PREDICTOR VARIABLES WITH
TOTAL ESSAY SCORE AS CRITERION

	Restriction	R_u^2	R_r^2	df	F-Ratio	Probab- ility
X(4)	Written Exp. Vocabulary	.576	.563	1/141	4.33	0.039
X(5)	Sentence Sensitivity	.576	.534	1/141	13.82	<0.001
X(6)	Effectiveness	.576	.566	1/141	3.35	0.069
X(7)	Mechanics	.576	.543	1/141	10.98	0.001
X(6)+X(7)	English Expression	.576	.534	2/141	6.70	0.001
X(8)	STEP: Writing	.576	.564	1/141	3.93	0.049
X(9)	Vocabulary	.576	.564	1/141	3.80	0.053
X(10)	Comprehension	.576	.576	1/141	0.00	1.000
X(9)+X(10)	Reading Comprehension	.576	.576	2/141	2.00	0.137
X(11)	Sex	.576	.576	1/141	0.12	0.721
X(12)	Intelligence	.576	.576	1/141	0.06	0.796

REGRESSION EQUATION FOR UNRESTRICTED MODEL

$$\begin{aligned}
 Y = & 0.096X(4) + 0.227X(5) - 0.237X(6) + 0.217X(7) \\
 & + 0.148X(8) + 0.166X(9) + 0.000X(10) - 0.259X(11) \\
 & - 0.014X(13) - 1.251
 \end{aligned}$$

*Level of significance = 0.05.

There is little difference between Tables XIII and XVII. The variables (4, 5, 7, 6 and 7, 8, 9) shown to be making a significant contribution in Table XIII continue to contribute significantly even in the presence of sex and intelligence. Hypothesis 6a is supported in that the unique contribution of the Sentence Sensitivity Test to the total essay score variance remains significant in the presence of all other variables including sex and intelligence. Hypothesis 6b on the unique contribution of the Written Expressional Vocabulary Test is also supported.

A similar result occurs when Tables XVI and XVIII, showing the contribution of all predictor variables to the prediction of scores on a two-essay criterion, are compared. Without including the sex and intelligence variables, all variables in Table XVI combine to predict 54.6% of the variance of composition scores, but with the addition of sex and intelligence variables, as in Table XVIII, 54.8% of the variance is accounted for. The increase in prediction is negligible. The variables making a significant contribution (1, 5, 7, 6 and 7) in Table XVI continue to contribute significantly in the presence of sex and intelligence, while sex and intelligence do not add to the prediction in the presence of the other variables. Hypothesis 6c is supported in that the unique contribution of the twenty-minute paragraph to the two-essay score variance is significant in the presence of all variables, including sex and intelligence.

TABLE XVIII

CONTRIBUTION OF PREDICTOR VARIABLES WITH
TOTAL SCORE ON TWO ESSAYS AS CRITERION

	Restriction	R_u^2	R_r^2	df	F-Ratio	Probab- ility
X(1)	Paragraph	.548	.501	1/140	14.48	<0.001
X(4)	Written Exp. Vocabulary	.548	.544	1/140	1.05	0.305
X(5)	Sentence Sensitivity	.548	.523	1/140	7.55	0.006
X(6)	Effectiveness	.548	.540	1/140	2.44	0.120
X(7)	Mechanics	.548	.522	1/140	7.90	0.005
X(8)	STEP: Writing	.548	.539	1/140	2.87	0.092
X(6)+X(7)	English Expression	.548	.517	2/140	4.79	0.009
X(9)	Vocabulary	.548	.542	1/140	1.92	0.167
X(10)	Comprehension	.548	.547	1/140	1.44	0.704
X(9)+X(10)	Reading Comprehension	.548	.542	2/140	0.96	0.383
X(11)	Sex	.548	.547	1/140	0.17	0.674
X(12)	Intelligence	.548	.547	1/140	0.36	0.549

REGRESSION EQUATION FOR UNRESTRICTED MODEL

$$\begin{aligned}
 Y = & 0.620X(1) + 0.038X(4) + 0.135X(5) - 0.153X(6) \\
 & + 0.141X(7) + 0.099X(8) + 0.084X(10) - 0.016X(11) \\
 & - 0.024X(12) - 0.816
 \end{aligned}$$

*Level of significance = 0.05.

A summary of these results together with a discussion of their implications, is presented in the following chapter.

CHAPTER VIII

SUMMARY, CONCLUSIONS, AND SUGGESTIONS FOR FURTHER RESEARCH

I. SUMMARY

Review of the Study

A group of 151 tenth-grade students wrote on three different essay topics and took five objective tests of writing ability during a two-week period in February and March of 1969. Each of four raters assigned scores from one to four to each of the three essays. The total of twelve scores thus assigned became the criterion for validating the five objective tests, first individually and then in various combinations. Later, the total essay score on two topics became the criterion when the third topic (a twenty-minute paragraph) was added as a predictor and the variables were again examined individually and in combinations. Of the five objective tests used, three were selected from present standardized objective tests and two were constructed by the investigator. Finally, sex and intelligence were included as predictors and the unique contribution of each variable to the essay score variance was re-examined.

Findings

The main findings of the investigation may be stated as follows:

1. The total essay score (reading reliability of .89 and score reliability of .82) constitutes a reliable criterion of writing performance. Reading reliability ranges from .43 (reliability of one rater on Essay Three) to .89 (estimate of reliability of four raters on three essays) and generally increases with the number of essay topics and the number of raters.

2. All five objective tests have significant positive correlations with each of the three essays and with the total essay score on the three essays. All five tests appear to have acceptable validity. With the total score on three essays as the criterion, validity coefficients ranging from .569 to .622 were obtained for the five tests.

3. Written Expressional Vocabulary, and Sentence Sensitivity, the two tests constructed by the Investigator, correlated .589 and .618 with the three-essay criterion.

4. Composites of two different objective tests produced multiple correlations ranging from .639 to .702 with the three-essay criterion. The three combinations with the highest correlations (.700 or above) contained the investigator's Sentence Sensitivity Test.

5. Eight of the nine composites of three different objective tests produced multiple correlations above .710 with the three-essay criterion.

6. The combination of all five objective tests produced a multiple correlation of .758 with the three-essay criterion.

7. With a two-essay criterion, composites of two objective tests, produced validity coefficients ranging from .569 to .662. When a twenty-minute paragraph was added to the same ten two-test composites these correlations ranged from .665 to .721, with an average increase of .070 per correlation coefficient.

8. With a two-essay criterion, composites of three objective tests (without the paragraph) produced validity coefficients ranging from .621 to .687, but with the addition of the paragraph to the same test combinations the validity coefficients were increased to range from .690 to .731.

9. Eight of the possible ten two-test combinations plus the paragraph produced higher correlations than the highest correlation of the nine three-test combinations of objective tests.

10. The combination of all five objective tests plus the paragraph produced a multiple correlation of .739 with the two-essay criterion.

11. When measured against a three-essay criterion, variables 4, 5, 7, (6 and 7), 8, 9* each make a significant contribution to the prediction of composition scores, when measured in the presence of the other variables, 4 to 10. Variables 6, 9 and 10, 10 do not significantly add to the prediction of composition scores in the combination of variables in which they are measured. As hypothesized, the Written Expressional Vocabulary Test and the Sentence Sensitivity Test both make significant contributions to the prediction of written composition scores.

*The tests represented by these numbers are listed on page 132.

12. When measured against a two-essay criterion, variables 1, 5, 6 and 7, 7 each contribute significantly in the presence of the other variables, while variables 4, 6, 8, 9, 10, 9 and 10 do not make a unique and significant contribution to the prediction of composition scores when each is measured in the presence of the other variables.

13. As hypothesized, the twenty-minute paragraph does make a significant contribution to the prediction of composition scores when it is used in the presence of variables 4 to 10.

14. The contributions of the Written Expressional Vocabulary Test, the Sentence Sensitivity Test, and the paragraph to the prediction of composition scores on the three-essay criterion remained significant when sex and intelligence were added as predictor variables.

II. CONCLUSIONS AND IMPLICATIONS

On the basis of the data and in view of the specific findings listed above, several generalizations seem justified.

The total essay score is reasonably reliable and can serve as a sound criterion for validating the five objective tests used in this investigation. It is also significant that this reliable criterion measure was obtained by using a holistic method of rating, that is, judgements of each essay were made after a reading for total impression. If one can include three essays and have each essay

read by four different raters, the reading reliability of the total score could be expected to be approximately .89 and the score reliability .82. Such essay topics, however, must be selected, assigned, and graded with the appropriate attention given to minimizing the assignment variable, the writer variable, and the rater variable. While for purposes of research such expenditures of time, human resources, and money are perhaps justified, for the average composition teacher in a day-to-day teaching situation such controls are too cumbersome and too involved. For instance, grading alone would involve four teachers - teachers who already claim to have too much marking. In view of the impracticality of obtaining such a reliable measure of writing, the more extensive use of reliable objective tests, shown by the results of this study to be reasonably valid measures of writing, particularly if used in combination, is recommended.

When the five objective tests (the three selected from standardized objective tests, and the two constructed by the investigator) are evaluated against a reliable criterion of writing performance they prove to be reasonably valid, as is evidenced by the multiple correlation of .76. Individually, the validity coefficients range from .56 to .62. Since, generally, the different tests are designed to measure somewhat different skills believed to be concomitants of writing ability, even more important results were obtained by comparing the criterion with various combinations.

Composites of two objective tests produced multiple correlations ranging from .639 to .702, while composites of three objective tests produced multiple correlations ranging from .684 to .741. These findings give support to the validity of the objective tests used, in that the tests tend to rank the students in the same order that readers ranked them on the basis of reading three essays. Such tests with reasonable validity might well serve to supplement the teacher's task in teaching and testing written composition. There are undoubtedly many occasions in the school year when objective tests which are theoretically and statistically related to writing ability and which can be administered and scored with relative ease might prove very valuable to teachers; for example, in revealing weakness in students' writing skills, in helping to section students for instruction, and in determining final grades.

The combinations of the different two- and three-set groupings of objective tests with the twenty-minute paragraph proved to be more valid when measured against the two-essay criterion than either test alone. In other words, the correlation of the paragraph with the two-essay criterion was .613, and the correlations of the three-set groups of objective tests with the two-essay criterion ranged from .621 to .687; but when the paragraph was added to these three-set groupings of objective tests the correlations with the criterion ranged from .690 to .731. Such a finding has direct implications for the debate between the proponents of objective tests

and those who prefer the essay test. Perhaps the debate as outlined in Chapter I has too often been conducted on an "either-or" basis. It is certainly conceivable, on the basis of the above data, that combinations of objective tests (which measure accurately some skills involved in writing) with a twenty-minute paragraph (which measures directly, if somewhat less accurately) might be more effective than either type of test alone. That eight of the possible ten two-test combinations plus the paragraph produced higher correlations than the highest correlation of the nine three-test combinations of objective tests (See Table VIII) makes this implication more inviting. The fact that the twenty-minute paragraph compared favorably with objective tests and contributed to the validity of objective test combinations suggests that the paragraph rated holistically evaluates some facets of writing which present objective tests do not assess.

III. SUGGESTIONS FOR FURTHER RESEARCH

The study just described has answered some questions and raised others. It has shown that a reliable measure of writing performance can be obtained by using three essays and four raters. However, the selection of a particular combination of essays for practical use will undoubtedly have to balance relative increase in reliability with the increase in student-writing and teacher-grading time accompanying this increase. At what point would the addition of further essays and more raters lead to a negligible increase in

reliability - especially viewed against increased student and teacher time?

This study has shown that indirect measurement of writing skills by means of objective tests has reasonably high validity and that particular combinations of objective tests can increase this validity. But it raises the question of what would be the effect of the addition of other objective tests. Would the construction and addition of other objective tests, based upon elements of the concept of writing ability and the writing process developed in Chapter II, account for more of the criterion score variance?

The study has shown that the two tests constructed by the investigator (Sentence Sensitivity and Written Expressional Vocabulary) are valid measures of selected writing skills, but no attempt was made to determine the specific contribution of the seven subtests of the first and the nine subtests of the latter. The effect of each subtest could in itself be investigated in detail with a larger number of test items.

The finding that a twenty-minute paragraph contributes to the validity of all objective-test combinations raises the following questions: What facets of writing are evaluated by the twenty-minute paragraph that are not evaluated by the objective tests? What would be the validity of the combination of two such paragraphs

and one objective test? Would different types of paragraphs make different contributions?

This study dealt only with a relatively small number of grade ten students. A large-scale experiment involving other grades and perhaps an entire provincial student population should be carried out to further determine the validity of objective tests of writing ability.

BIBLIOGRAPHY

BIBLIOGRAPHY

- Allen, D. A. "Review of STEP-Writing." 595-597, in Buross, O. D. The Sixth Mental Measurements Yearbook. Highland Park, New Jersey: The Gryphon Press, 1965.
- Anderson, C. C. "The New STEP Essay as a Measure of Composition Ability," Educational and Psychological Measurement, XX (Spring, 1960), 95-102
- Armstrong, R. D. An Objective Measure of the Quality of the Written Composition of Fifth-Grade Pupils. University of California, Berkeley, Ed. D., 1965 (University Microfilms, Inc., Ann Arbor, Michigan).
- Bateman, D. and F. Zidonis, The Effect of a Study of Transformational Grammar on the Writing of Ninth and Tenth Graders. National Council of Teachers of English, Research Report No. 6, Champaign, Illinois: National Council of Teachers of English, 1966.
- Black, D. B. "A Note on the Use in Alberta of Sequential Tests of Educational Progress: Essay Test," Alberta Journal of Educational Research, IV (September, 1958), 172-180.
- Booth, W. C. "The Rhetorical Stance," 181-189, in Tate, G. and E.P.J. Corbett, Teaching Freshman Composition. New York: Oxford University Press, 1967.
- Bottenberg, R. A. and J. H. Ward, Jr. Applied Multiple Linear Regression. Clearing House for Federal Scientific and Technical Information, U.S. Department of Commerce, Technical Documentary Report PRL-TDR-63-6. Washington: Government Printing Office, 1963.
- Braddock, R. Lloyd-Jones, R. and L. Schoer, Research in Written Composition. Champaign, Illinois: National Council of Teachers of English, 1963.
- Burgess, J. "Sentence by Sentence," College Composition and Communication. XIV (December, 1963), 257-262.
- Buxton, E. W. "An Experiment to Test the Effects of Writing Frequency and Guided Practice Upon Students' Skill in Written Composition," Alberta Journal of Educational Research, V (June, 1959), 91-99.

- Cain, T. H. Common Sense about Writing. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1967.
- Carroll, J. B. "Psycholinguistics and the Teaching of English Composition," College Composition and Communication. VII (December, 1956), 188-193.
- Carroll, J. B. "The Critical Need in the Study of Language," 23-26 in Linguistic, Composing, and Verbal Learning, (Papers from the 1962 Conference on College Composition and Communication).
- Chomsky, N. "A Review of B.F. Skinner's Verbal Behavior," 547-578, in Fodor, J. A. and J. J. Katz, The Structure of Language. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1964.
- Christensen, F. Notes Toward A New Rhetoric. New York: Harper and Row, Publishers, 1967.
- Christensen, F. "The Problem of Defining a Mature Style," English Journal. LVII (April, 1968), 572-579.
- Christiansen, M. A. "An Experimental Study in Composition: Extensive Writing vs. Some Writing Plus Reading, University of Kansas Bulletin of Education, XIX (May, 1965), 124-131.
- Commission on English (Report of), Freedom and Discipline in English, New York: College Entrance Examination Board, 1965.
- Conlin, D. A. Grammar for Written English. Boston: Houghton Mifflin Company, 1961.
- Cordor, J. W. Rhetoric: A Text Reader on Language and Its Uses. New York: Random House, 1965.
- Coward, A. F. "A Comparison of Two Methods of Grading English Compositions," The Journal of Educational Research 46 (1952), 5-18.
- Cowley, M. Writers at Work: The Paris Review Interviews. New York: The Viking Press, Inc., 1967.
- Davis, F. B. "Fundamental Factors of Comprehension in Reading," Psychometrika. IX (September, 1944), 185-197.
- Diederich, P. B. "The Measurement of Skill in Writing," School Review, LIV (December, 1946), 585-587.

- Emig, J. "On Teaching Composition: Some Hypotheses as Definitions," Research in the Teaching of English. I (Fall, 1967), 127-135.
- Enochs, J. B. Measuring the Ability to Write. Chicago: University of Chicago, 1948.
- Erskine, J. "The Craft of Writing." quoted in F. Christensen, Notes Toward A New Rhetoric. New York: Harper and Row, Publishers, 1967.
- Fleming, W. G. "Review of Reading Comprehension: Cooperative English Tests," 1084-1086, in Buros, O. K. The Sixth Mental Measurement Yearbook. Highland Park, New Jersey: The Gryphon Press, 1965.
- Follman, J. C. and J. A. Anderson, "An Investigation of the Reliability of Five Procedures for Grading English Themes," Research in the Teaching of English. I (Fall, 1967), 190-200.
- Postvedt, D. R. "Criteria for the Evaluation of High-School English Composition," The Journal of Educational Research. 59(November, 1965), 108-112.
- Fowler, M. E. Teaching Language, Composition, and Literature. New York: McGraw-Hill Book Company, 1965.
- French, J. W. Schools of Thought in Judging Excellence of English Themes, Princeton, New Jersey: Educational Testing Service, 1965.
- Frogner, E. "Problems of Sentence Structure in Pupils' Themes," The English Journal. XXII (November, 1933), 742-749.
- Gale, C. Building an Effective Vocabulary. New York: Barron's Educational Series, Inc., 1966.
- Godshalk, F. I. "A Reply to Critics," Research in the Teaching of English. I (Spring, 1967), 84-88.
- Godshalk, F. I., Swineford, F. and W.E. Coffman, The Measurement of Writing Ability. New York: College Entrance Examination Board, 1966.
- Ghiselin, B. The Creative Process. New York: The New American Library of World Literature, Inc., 1961.
- Greene, H. A. "English-Language, Grammar, and Composition," Encyclopedia of Educational Research, (New York: MacMillan Co., 1950), 383-396.

- Guth, H. P. English Today and Tomorrow. New Jersey: Prentice Hall, Inc., 1964.
- Hamalian, L. and E. L. Volpe, Grammar in Context. New York: G. P. Putnam's Sons, 1960.
- Heys, F. "A Theme A Week Assumption: A Report of an Experiment," English Journal LI (May, 1962), 320-322.
- Hinkelman, R. "Relationship of Reading Ability to Elementary School Achievement," Educational Administration and Supervision, XLII (February, 1956), 65-67.
- Huddleston, E. M. "Measurement of Writing Ability at the College-Entrance Level: Objective vs. Subjective Testing Techniques," Journal of Experimental Education. XXII (March, 1954), 165-213.
- Hunt, K. W. Grammatical Structures Written at Three Grade Levels. NCTE Research Report No. 3, Champaign, Illinois: National Council of Teachers of English, 1965.
- Hunting, R. "Recent Studies of Writing Frequency," Research in the Teaching of English. I (Spring, 1967), 29-40.
- Jenkins, J. J. "Mediation Processes and the Acquisition of Linguistic Structure," quoted in Bateman, D. and F. Zidonis, The Effect of A Study of Transformational Grammar on the Writing of Ninth and Tenth Graders. NCTE Research Report No. 6, Illinois, National Council of Teachers of English, 1966.
- Jewett, A. Improving English Composition. Washington: National Education Association, 1965.
- Johnson, M. F. "On the Relationship Between Sentence Structure and the Latency in Generating the Sentence," and "The Psychological Reality of Phrasestructure Rules." quoted in Bateman, D. and F. Zidonis, The Effect of A Study of Transformational Grammar on the Writing of Ninth and Tenth Graders. NCTE Research Report No. 6, Illinois: National Council of Teachers of English, 1966.
- Jones, E. A. "An Investigation of the Relationship Between Written Composition and Reading Ability and Reading Habits," Unpublished Master's Thesis, University of Alberta, 1966.
- Katz, J. J. "Mentalism in Linguistics," Language. 40 (April-June, 1964).

- Kincaid, G. L. "Some Factors Affecting Variations in the Quality of Students' Writing," 83-94, in Braddock, R. Lloyd-Jones, R. and L. Schoer, Research in Written Composition, Illinois, National Council of Teachers of English, 1963.
- Kitzhaber, A. R. Themes, Theories, and Therapy: The Teaching of Writing in College. New York: McGraw-Hill Book Company, Inc., 1963.
- LaBrant, L. L. "Studies of Certain Language Developments of Children in Grades Four to Twelve." Genetic Psychology Monographs. XIV (November, 1933), 387-494.
- Lampard, D. M. "An Examination of the Construction and Reconstruction of Meaning as Seen in the Written Responses Made to Questions about What Has Been Read by Students in the High Schools of Alberta," Unpublished Ph.D. Dissertation, University of Oregon, 1966.
- Loban, W. D. The Language of Elementary School Children. NCTE Research Report No. 1, Illinois: National Council of Teachers of English, 1963.
- Loban, W., Ryan, M. and J. R. Squire, Teaching Language and Literature. New York: Harcourt, Brace and World, Inc., 1961.
- Lowe, F. "Writers on Learning to Write," English Journal. 53 (October, 1963), 488-495.
- Lecky, E. "Rhetoric, Semantics, and Composition," 5-15, in R. M. Gorrell, (ed.), Rhetoric: Theories for Application. Illinois: National Council of Teachers of English, 1967.
- Lyman, R. L. Summary of Investigations Relating to Grammar, Language, and Composition. Chicago, Illinois: The University of Chicago, 1929.
- Martin, H. C. and R. M. Ohman, The Logic and Rhetoric of Exposition. New York: Holt, Rinehart, and Winston, Inc., 1963.
- McCullough, C. M. and J. C. Flanagan, "The Validity of the Machine-Scorable Cooperative English Tests," Journal of Experimental Education. VII (March, 1939), 229-234.
- Myers, A. E. McConville, C. B. and W. E. Coffman, "Simplex Structure in the Grading of Essays," Educational and Psychological Measurement. XXVI (1966), 41-53.

- Newsome, V. "Expansions and Transformations to Improve Sentences," English Journal. 53 (1964), 327-335.
- Palmer, O. "The Objective Testing of English Composition," English Journal. L (May, 1961), 314-320.
- Perrin, P. G. Writer's Guide and Index to English. Chicago: Scott, Foresman and Company, 1965.
- Petty, W. T., Herold, C. P. and E. Stoll. The State of Knowledge about the Teaching of Vocabulary. Champaign, Illinois: National Council of Teachers of English, 1968.
- Pitt, J. and R. E. Leavenworth. Logic For Argument. New York: Random House, 1968.
- Porter, D. "The Behavioral Repertoire of Writing," 14-17, in Linguistics, Composing and Verbal Learning (Papers from the 1962 Conference on College Composition and Communication).
- Potter, R. R. "Sentence Structure and Prose Quality: an Exploratory Study," Research in the Teaching of English. I (Spring, 1967), 17-27.
- Rohman, D. G. "Pre-Writing: The Stage of Discovery in the Writing Process," College Composition and Communication. 16 (May, 1965), 106-112.
- Ross, C. C. and J. C. Stanley. Measurement in Today's Schools. New York: Prentice Hall, 1954.
- Sherwood, J. C. "English Expression: Cooperative English Tests (1960 Revision)," 556-558 in Buros, The Sixth Mental Measurements Yearbook. Highland Park, New Jersey: Gryphon Press, 1965.
- Skinner, B. F. Verbal Behavior. New York: Appleton-Century-Crofts, Inc., 1957.
- Smith, W. F. and R. D. Liedlich, From Thought to Theme. New York: Harcourt, Brace and World, Inc., 1968.
- Sochor, E. E. "The Nature of Critical Reading," 3-14, in Critical Reading: An Introduction. Illinois, National Council of Teachers of English, 1959.

- Steinmann, Jr. M. "A Conceptual Review," Research in the Teaching of English. 1 (Spring, 1967), 74-84.
- Steinmann, Jr. M. "Rhetorical Research," 16-32 in M. Steinmann Jr., New Rhetorics. New York: Charles Scribner's Sons, 1967.
- Thorndike, E. L. "Reading as Reasoning: A Study of Mistakes in Paragraph Reading," Journal of Educational Psychology. VIII (June, 1917), 323-332.
- Traxler, A. E. and H. A. Anderson, "The Reliability of an Essay Test in English," School Review. XLIII (September, 1935), 534-539.
- Waterston, E. and M. Beattie, Composition for Canadian Universities. Toronto: The MacMillan Company of Canada Limited, 1964.
- Watt, A. F. The Language and Mental Development of Children. London: D. C. Heath and Company, 1944.
- Willis, H. Structural Grammar and Composition. New York: Holt, Rinehart and Winston, Inc., 1967.
- Wilson, G. and R. J. Lacompage. "Developments in Composition," The Bulletin of the National Association of Secondary School Principals, 51 (April, 1967), 60-66.
- Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill Company, 1962.

TESTS, MANUALS, AND TECHNICAL REPORTS

- Cooperative Test Division, Cooperative Sequential Tests of Educational Progress: Essay Test, Form 2A. Princeton, New Jersey: Educational Testing Service, 1957.
- Cooperative Test Division, Handbook for Essay Test: Level 2. Princeton, New Jersey: Educational Testing Service, 1957.
- Cooperative Test Division, Sequential Tests of Educational Progress: Writing, Form 2B. Princeton, New Jersey: Educational Testing Service, 1957.

Cooperative Test Division, Manual for Interpreting Scores: Writing. Princeton, New Jersey: Educational Testing Service, 1957.

Cooperative Test Division, Technical Report: Writing. Princeton, New Jersey: Educational Testing Service, 1957.

Cooperative Test Division, Cooperative English Tests: Reading Comprehension, Form 2A. Princeton, New Jersey: Educational Testing Service, 1960.

Cooperative Test Division, Cooperative English Tests: English Expression: Form 2A. Princeton, New Jersey: Educational Testing Service, 1960.

Cooperative Test Division, Cooperative English Tests: Manual for Interpreting Scores. Princeton, New Jersey: Educational Testing Service, 1960.

APPENDICES

APPENDIX A

ESSAY TESTS

Essay One - STEP Essay: Form 2A

Essay Two - STEP Essay: Form 2D

051

SCORE

NAME	(PRINT)	LAST	FIRST	MIDDLE
AGE	YEARS	MONTHS	GRADE OR CLASS	
SCHOOL				
TODAY'S DATE		MONTH	DAY	YEAR
1.		2.		3.

DO NOT WRITE HERE

Cooperative
Sequential Tests of Educational Progress
Essay Test

General Directions

The purpose of this test is to find out how well you can write an essay. First you will read a short passage which will tell you what you are to write about. Then you should plan what you want to say and the order in which you want to say it. As soon as you have finished planning, you should begin to write. You will have about thirty minutes for writing.

Here are a few suggestions which will help you to do your best on the test:

1. Start planning your paper as soon as you know what you are to do. You may use the space underneath the reading passage or the back of your booklet for making notes about your plans.
2. While spelling and punctuation will be considered in grading your paper, what you have to say and how well you say it will be more important. Therefore, you should spend most of your time getting your ideas down in a clear, well-organized form. Watch your handwriting, too; your paper cannot be marked if no one can read it.
3. Save a little time at the end to check your paper and make any needed changes. Since there will not be time to copy what you write, make your changes neatly by writing between the lines.

Do not turn the page until you are told to do so.

Teen-agers often find their situation confusing, because they are sometimes treated as adults, sometimes as children. When should teen-agers be treated as adults? You might consider some of the following questions in making your decision:

- (a) Are you mature when you reach a certain age? (The government considers eighteen a satisfactory age for military service.)
- (b) Are you mature when you graduate from high school? (The high school diploma indicates that you are ready for work or for college.)
- (c) Are you mature when you have a regular job and are self-supporting? (Many high school students have regular jobs.)
- (d) Are you mature when you can accept responsibility? (Some high school freshmen are more conscientious and dependable than seniors.)

When do *you* think teen-agers should be treated as adults? Give reasons for your opinion, and support it by specific suggestions or examples.

ESSAY TEST

GENERAL DIRECTIONS

The purpose of this test is to find out how well you can write an essay. First you will read a short passage which will tell you what you are to write about. Then you should plan what you have to say and the order in which you want to say it. As soon as you have finished planning, you should begin to write. You will have about thirty minutes for writing.

Here are a few suggestions which will help you to do your best on the test:

1. Start planning your paper as soon as you know what you are to do.
2. While spelling and punctuation will be considered in grading your paper, what you have to say and how well you say it will be more important. Therefore, you should spend most of your time getting your ideas down in a clear, well-organized form. Watch your handwriting, too; your paper cannot be marked if no one can read it.
3. Save a little time at the end to check your paper and make any needed changes. Since there will not be time to copy what you write, make your changes neatly by writing between the lines.

TOPIC

Last year's student-body president came back to visit his high school just before Christmas vacation. He had just completed a term at the university. At a regular assembly he was invited to sit on the platform and say a few words. This was his message:

"I've been down on the university campus since September and it's great. High school's all right, but college is the place for fun - fraternities, and of course sororities for the girls. Football, with rallies before and after the big games; dances and parties, and of course I got involved right away in campus politics. There's a full program of studies for everyone, too. You study there harder than any studying you ever did in high school. But it's really the life. I'd like to leave this idea with you. Everyone of you ought to go to college. I mean it - I'd like to see all high school graduates of every high school go on to college, and I know that all of you, if you get to college, you'll think it's great."

How well can you analyze the speaker from what he says? Tell what you think about his temperament, his character, his ability to think. Limit your character analysis to traits which you can interpret from the direct quotation. Explain what clues led you to draw your conclusions.

APPENDIX B

Written Expressional Vocabulary Test

Sentence Sensitivity Test

WRITTEN EXPRESSIONAL VOCABULARY TEST

This test deals with several important aspects of effective word usage:

- (1) your variety of expression, or your ability to express particular concepts through the use of different words;
- (2) your precision of expression, or the degree of discrimination you exercise in selecting one of a number of near synonymous terms as the most appropriate to represent the meaning to be conveyed; and
- (3) the appropriateness of your word usage.

Directions

This is a 30-minute test with 9 sections. The directions for each section are printed at the beginning of the section. When you are told to begin, turn this page and read the directions for the first section. Then go on immediately to answer the questions for that section. Mark your answers in the space provided in the test booklet. After you have finished the first section go on to the other sections. Read all directions and follow them carefully. Work at a steady pace. If you come to an item you cannot answer, leave it and try it later if you have time. Your score will be the total number of items you have correct.

Do not turn this page until you are told to do so.

For each of the words underlined in the following phrases, write three synonyms in the blanks. (Synonyms are words having a meaning that is the same or nearly the same as the underlined word.)

Example: a dangerous position	<u>hazardous</u>	<u>perilous</u>	<u>precarious</u>
1. to <u>predict</u> the future	_____	_____	_____
2. <u>cold</u> areas of the north	_____	_____	_____
3. keeping <u>careful</u> watch	_____	_____	_____
4. to <u>consider</u> the evidence	_____	_____	_____
5. <u>turn away</u> from one's course	_____	_____	_____

Supply two synonyms for the word wild in each of the following sentences. A word used in one answer will not be accepted as an answer to another question.

1. He was captured by the wild tribes. _____
2. What are the wild waves saying? _____
3. He saw a wild rabbit. _____
4. He is full of wild notions. _____
5. The baseball pitcher is wild today. _____

When you take action and movement and make it highly specific you flash a vivid picture on the minds of your readers. Take the sentence, "'No,' the girl said, and left the room." The idea that these words convey is both simple and clear. It is also absolutely colourless, for say and leave are among the most general verbs in the language.

Suggest three words you might use in place of the word said in the following situations:

1. Suppose "the girl" left speaking loudly and displaying angry passion or vehemence.

"'No,' the girl _____, and left the room."

2. Suppose "the girl" left complaining in a bad-tempered way with a low voice

"'No,' the girl _____, and left the room."

3. Suppose "the girl" left speaking tearfully in tones of sadness, sorrow, or grief.

"'No,' the girl _____, and left the room."

Suggest three words you might use in place of the word left in the following situations:

4. Suppose "the girl" left the room stealthily and trying to be unnoticed.

"'No,' the girl said, and _____ the room."

5. Suppose "the girl" left the room quickly and in a state of great excitement.

"'No,' the girl said, and _____ the room."

Each question in this group consists of a sentence from which two words have been omitted. Below each incomplete sentence are five pairs of words. You must choose one of these groups to complete the sentence and make it a true statement. In the space provided, write the letter of the set of words that best completes the thought of the sentence group.

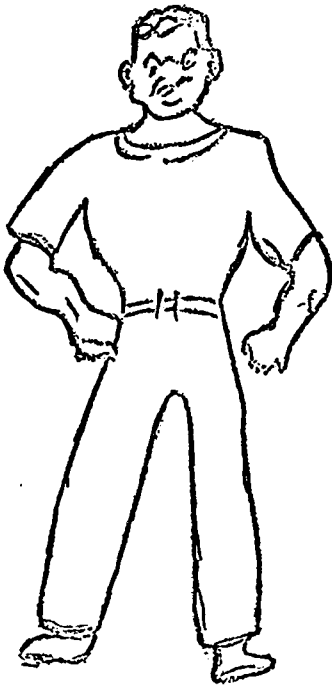
1. A _____ response is one that is made with _____.
(a) stupid, fear (b) speedy, alacrity (c) sure, slowness (d) harmful, grimaces (e) pleasant, surmise _____
2. Into the limited space given him a headline writer must compress the _____ of the news and he must do it without _____.
(a) bias, apology (b) magnitude, distortion (c) totality, hedging (d) synopsis, suggestions (e) gist, ambiguity _____
3. To limit a press is to insult a nation; to _____ reading of certain books is to declare the _____ to be either fools or slaves.
(a) encourage, readers (b) allow, authors (c) censure, youth (d) restrict, librarians (e) prohibit, inhabitants _____
4. _____ minds generally _____ everything that passes their understanding.
(a) Great, disregard (b) Mediocre, condemn (c) First-rate, study (d) Simple, admire (e) Educated, read _____
5. The world will never know of my life if it should write and read a hundred _____. The main facts of it are known, and are likely to be known, to myself alone, of all _____ men.
(a) biographies, created (b) biographies, unknown (c) stories, famous (d) articles, injured (e) autobiographies, fortunate _____

A word has been omitted in each of the sentences below. Decide which of the suggested words the author probably used and circle the appropriate letter. Let the context be your guide.

1. The story was told, retold, and _____ with each telling.
(a) adorned (b) decorated (c) embellished (d) enriched
2. Many public services were _____ because of the emergency.
(a) diminished (b) reduced (c) curtailed
3. I should not feel confident in venturing on a journey in a foreign country without a(n) _____.
(a) associate (b) chum (c) companion (d) friend
4. As the waves rose and the ship tossed, many of the passengers felt _____.
(a) lethargic (b) subdued (c) tremulous (d) queasy
5. The girls _____ excitedly when they heard plans for the party.
(a) chattered (b) talked (c) babbled
6. The only way to _____ arguments is to discuss long in advance with the entire group just what the problems to be faced are and what course of action should be assumed.
(a) settle (b) forestall (c) judge (d) appease

Indicate, by placing the appropriate letter, A, B, or C in front of each of the following 20 words, to which man the words could best apply.

A



B



C



___ belligerent

___ corpulent

___ dapper

___ despondent

___ jubilant

___ dishevelled

___ emaciated

___ gaunt

___ rotund

___ seedy

___ morose

___ melancholy

___ truculent

___ pugnacious

___ genial

___ hearty

___ obese

___ portly

___ robust

___ vigorous

In each of the following sentences, two or more words appear in parentheses. Find another word - one only - that could be substituted for the parenthetical expression, and write it in the space provided.

1. Do not tell stories in company; there is nothing more (tiresome and boring) and disagreeable. _____
2. That movie was really (one which made you think deeply). _____
3. I have simply clothed my thoughts in what appeared to me the most obvious and (especially suitable) language. _____
4. I was a very (small and thin) lad. _____
5. The explosion was (one of tremendous power and of frightening impact). _____

From each of the following word groups select the word having the most favorable connotation and write the appropriate letter of that word in the space provided. (Connotation means the suggestive or extra meaning of a word rather than its core or denotational meaning.)

1. (a) slim (b) thin (c) skinny _____
2. (a) dictator (b) strong man (c) absolute ruler _____
3. (a) new (b) new-fangled (c) up-to-date _____
4. (a) outmoded (b) time-tested (c) old _____
5. (a) girl (b) wench (c) lass _____
6. (a) felicity (b) bliss (c) happiness _____
7. (a) polite (b) courteous (c) civil _____
8. (a) leader of the people (b) party leader (c) demagogue _____

In the sentences below, the words or phrases in parentheses are clichés that the writer did not use. In the space provided, supply words that you think the writer might have used.

1. I would give you some violets, but they are (dead as a doornail).

2. Agreement is an altogether tiresome element in (yackety-yack).

3. A woman near me was (in a vale of tears), but when I looked at her, I saw that her face was transfigured.

4. He turned out countless articles and sketches that gave him pleasure only because they contained a (teeny weeny bit) of what was in his mind.

5. Her knees were shaky and she felt (dog) tired as she pushed her way through the crowd to the information desk.

SENTENCE SENSITIVITY TEST

(This test deals with several important aspects of making effective sentences.)

Directions

This is a 30-minute test with 7 sections. The directions for each section are printed at the beginning of the section. When you are told to begin, turn this page and read the directions for the first section. Then go on immediately to answer the questions for that section. Mark your answers in the space provided in the test booklet. After you have finished the first section go on to the other sections. Read all directions and follow them carefully. Work at a steady pace. If you come to an item you cannot answer, leave it and try it later if you have time. Your score will be the total number of items you have correct.

Do not turn this page until you are told to do so.

Each of the ten pairs of sentences below shows two ways of expressing an idea. Select the sentence that you feel gives greater emphasis to the idea the writer wants expressed. (Emphasis, in writing sentences, comes from the use of strong and distinctive words, from effective punctuation, and from the position or ordering of words and parts of the statement so that attention is directed unmistakably to the idea you want to stress.) For each pair of sentences circle the letter of the sentence you select.

1. a. There are several employees who have been overstaying their coffee break.
b. Several employees have been overstaying their coffee break.
2. a. We are distracted by the way you tap your pencil.
b. It is distracting to us the way you tap your pencil.
3. a. Both starting halfbacks failed the final examination.
b. The final examination was failed by both starting halfbacks.
4. a. The instructor said he would grade the papers in two or three days.
b. The instructor said that the papers would be graded by him in two or three days.
5. a. If they do not build another school soon, this one will be hopelessly overcrowded.
b. If another school is not built soon, this one will be hopelessly overcrowded.
6. a. He was accused of cheating and was expelled from school by the Disciplinary Committee yesterday afternoon at a meeting.
b. He was accused of cheating and at a meeting of the Disciplinary Committee yesterday afternoon, he was expelled from school.
7. a. The favorite thrilled the crowd by a magnificent stretch run in which he overtook six horses and won by a nose.
b. In a magnificent stretch run the favorite overtook six horses and won by a nose, thrilling the crowd.
8. a. This hawk's small and weak talons confine its attention to little animals.
b. This hawk's talons, small and weak, confine its attention to little animals.
9. a. In the course of a lifetime of voyaging he went to China, India, all parts of Africa, and even the arctic.
b. In the course of a lifetime of voyaging he went to China, to India, to all parts of Africa, and even to the arctic.

10. a. Many people read history to raise their self-esteem.
b. There are many people who read history to raise their self-esteem.

0

In the following sentences, underline any word or words that are not necessary to convey the full meaning.

Example: These things, though of a useful nature, were not what he wanted.

1. Mr. Smith died before he had a chance to write his own autobiography.
2. The recorded cases of violence are countless in number.
3. It was during this time that the greatest number of cases were judged.
4. I have a friend of mine in the business game who would like to read this book written by Dr. Crocker.
5. This is a youth who wandered from the group.
6. There are several explanations in the book which throw light on the behavior of young adolescents.
7. Every thinking person these days seems inclined to agree with the conception that the world has gone mad.
8. To my surprise the damage was not so bad as I had expected it to be.
9. The following statistics serve to give a good idea of the effects of tobacco.

Combine the following pairs of short sentences into one well-formed sentence. You may add or delete some words if you wish, but remember to retain the original idea of each sentence.

Example: a. We did not hurry.
b. We knew that the bus was late.

Combined: Knowing that the bus was late, we did not hurry.

1. a. I am going to the rodeo.
b. My sister is going too.
-

2. a. She bought a new dress.
b. She also bought a purse and gloves to match.
-

3. a. I have a friend who is a certified public accountant.
b. He works in a large city bank, and has urged me to take up accounting.
-
-

4. a. The salesman was angered by our refusal to buy the car.
b. He became very insulting.
-

5. a. His argument was easily proved false.
b. It was that girls are smarter than boys.
-

6. a. Dad insisted on my taking fourth year English.
b. He knew the value of literature and composition.
-

7. a. We are planning a reunion in June.
b. It will honor last year's graduating class.
-

8. a. He ate blueberry pie.
b. He came home from school.
-

9. a. Who is that girl?
b. I don't know.
-

10. a. The sailors tried to rescue the captain.
b. They soon had the lifeboats over the side.
-
-

Combine the following groups of short sentences into one well-formed, mature sentence. You may add or delete some words if you wish, but remember to retain the original idea of each sentence.

1. a. I'm looking for my little sister.
b. She is fat.
c. She has blue eyes.
d. She was eating an ice cream cone.
-
-

2. a. The road was narrow and wet.
b. Tom had to drive carefully.
c. He arrived at Joan's house at ten.
-
-

3. a. The war was no fun for the front-line soldiers.
b. They lived with danger.
c. They lived with hardships.
d. They lived with a total lack of pleasure.
-
-

4. a. Grandfather inched up the stairs.
b. He grasped the handrail with one hand.
c. He leaned on his cane with the other.
d. He firmly planted both feet on each step.
-
-

5. a. I offered my help.
b. I offered it with reluctance.
c. I soon found I had good reason to be reluctant.
d. I was spattered with mud from head to toe.
-
-

In each group below, the first sentence was written by a famous author. In the margin, circle the letter of the sentence that appears to you to be the best "next" sentence. (Possible considerations include: which most advanced the thought? maintains suspense? uses words economically? fits the rhythm? creates the best image? maintains the mood? avoids extraneous matters?)

1. He reproached himself bitterly for his behavior that evening.
 - a. But why reproach yourself?
 - b. He was absolutely certain, as sure as he was that Sunday was Sunday, that she would never see him again?
 - c. Of course, it wasn't his fault.
 - d. Was there any reason in the world for her to want to see him again?
 - e. Why had he given her the alternative that she must dine with him or else never see him again?

a b c d e
2. A degenerate nobleman, or one that is proud of his birth, is like a turnip.
 - a. A tough skin hides the poor food beneath.
 - b. Both improve when well-seasoned.
 - c. In the market place, neither can command a high price.
 - d. There is nothing good of him but that which is underground.
 - e. They are both vegetables.

a b c d e
3. There are no necessary evils in government.
 - a. After all, people make up a government.
 - b. Good may exist in a legislative body.
 - c. Governments are composed of many branches and each has to be judged separately.
 - d. Its evils exist only in its abuses.
 - e. Though religious people may do evil, the religion itself is still good.

a b c d e
4. More important than winning the election is governing the nation.
 - a. It is important that the spoils do not go to the victor.
 - b. Nations are not built by elections.
 - c. That is the test of a political party - the avid, final test.
 - d. This governing has to be honest, thoughtful, and courageous.
 - e. This is why Canada is a great country.

a b c d e
5. .. It is a comfortable feeling to know that you stand on your own ground.
 - a. By ground I mean not necessarily a bit of earth, but your own home, your own flat, your own small room.
 - b. Can there be a better feeling than knowing that you are secure on what belong to you.
 - c. Even a dog barks with greater surety when he is on his master's property.
 - d. Ground is solid.
 - e. Land is about the only thing that can't fly away.

a b c d e

In each of the following sets of four sentences, one is a better formed, more mature sentence than the other three. Circle the letter of the best sentence in each group.

1. a. The banker agreed to lend mother the money, which made her very happy.
b. The banker agreed to lend mother the money; this made her very happy.
c. The banker agreed to lend mother the money; which made her very happy.
d. The banker agreed to lend mother the money, this making her very happy.
2. a. Byron not receiving a college scholarship disappointed his parents.
b. Being that Byron did not receive a college scholarship disappointed his parents.
c. Because Byron did not receive a college scholarship was the reason he disappointed his parents.
d. That Byron did not receive a college scholarship disappointed his parents.
3. a. Lincoln said that a house divided against itself could not stand, and this is certainly as true today as it was in 1858.
b. Lincoln said that a house divided against itself could not stand, which is certainly as true today as it was in 1858.
c. Lincoln said that a house divided against itself could not stand, an assertion which is certainly as true today as it was in 1858.
d. Lincoln said that a house divided against itself could not stand, and it is certainly as true today as it was in 1858.
4. a. His great courage and dogged determination makes him a good leader.
b. His great courage as well as his dogged determination make him a good leader.
c. His great courage and dogged determination make him a good leader.
d. His great courage added to his dogged determination make him a good leader.
5. a. Despite the poverty in southern Spain, they seem a fairly happy people.
b. Despite the poverty in southern Spain, they seem a fairly happy group of people.
c. Despite the poverty in southern Spain, happiness seems to be fairly common.
d. Despite the poverty in southern Spain, the people there seem fairly happy.

The following statements test your ability to arrange ideas in narrative and descriptive sentences. When the statements are placed in an appropriate order, one idea follows logically from another and the whole expression adds to a sequence that carries the reader through the subject. Decide on the arrangement that you think best conveys the appropriate order and indicate this order by arranging in the space at the right the letters of the expressions.

Example: a. still wiping his face

b. in a few minutes he came back

c. as if there were cobwebs on it

b a c

1. a. dragging its head through the weeds that hid the fence

b. I slowed still more

c. my shadow passing me

2. a. the petals fluttering with her movements

b. an orchid corsage pinned to her dress above her left shoulder

c. she sat surrounded by packages

3. a. the light flaming yellow white

b. she walked beside him down the dark path

c. dimly showing the damp, slippery, matted leaves

d. his left hand holding the cigarette lighter

4. a. he came out at once

b. his hands thrust deep into the pockets

c. wearing the yellow silk dressing gown

5. a. he walked with his head thrust forward

b. as if he had a slight curvature of the spine

c. so that his shoulders seemed raised and rounded

6. a. its shadow swinging too

b. then the head of the steamboat began to swing across the stream

c. travelling long before it across the water

7. a. bringing his right hand close to his hip

b. he turned quickly

c. so that the cape wound high and flat like a disk around him

APPENDIX C
RAW SCORES ON ALL TESTS

Variables for Raw Data

1. Identification Number
2. Essay One: STEP Essay, Form 2A
3. Essay Two: STEP Essay, Form 2D
4. Paragraph
5. Written Expressional Vocabulary
6. Sentence Sensitivity
7. Effectiveness
8. Mechanics
9. STEP Writing, Form 2B
10. Vocabulary
11. Comprehension
12. Sex: 0 if Female, 1 if Male
13. Intelligence

1	2	3	4	5	6	7	8	9	10	11	12	13
001	08	07	04	44	29	10	23	26	35	30	1	118
002	07	08	07	41	43	14	27	32	34	45	1	124
003	13	11	08	58	37	18	31	51	37	32	0	119
004	08	10	10	47	41	18	29	36	39	47	0	126
005	06	06	12	67	50	20	30	38	37	47	1	125
006	09	08	11	60	48	22	36	42	44	48	1	131
007	09	06	11	53	43	20	22	37	42	52	1	132
008	10	06	08	53	34	17	30	36	38	41	1	119
009	10	10	10	32	43	12	38	38	35	43	0	125
010	07	07	09	46	42	19	24	44	41	39	1	132
011	10	10	09	50	53	20	38	40	37	43	0	128
012	12	11	12	55	50	25	26	40	49	44	1	134
013	06	07	08	41	41	21	33	36	38	34	0	119
014	11	11	11	48	47	20	37	44	45	49	1	131
015	10	11	11	60	49	18	31	44	28	32	0	112
016	09	06	08	53	43	21	36	35	40	47	1	129
017	10	10	11	51	49	20	35	39	37	42	0	119
018	13	10	07	52	49	18	42	36	32	32	0	111
019	10	08	09	58	61	17	34	45	36	44	0	118
020	10	07	09	48	42	20	36	40	34	43	1	125
021	09	09	09	47	40	18	32	38	36	40	1	105
022	07	07	06	44	39	20	35	39	38	45	1	128
023	11	10	09	59	55	24	32	42	45	45	0	126
024	06	04	05	34	32	16	28	33	32	21	1	111
025	06	07	08	44	41	18	29	38	42	40	1	113
026	11	12	08	46	53	18	33	38	38	41	0	120
027	09	07	13	52	43	18	27	36	43	41	1	115
028	09	08	09	50	40	16	25	39	34	34	1	106
029	08	06	07	57	48	19	27	42	38	35	1	132
030	11	11	09	52	43	23	34	29	35	22	1	107
031	09	07	09	45	46	18	41	38	40	41	0	112
032	12	04	09	56	44	16	37	33	39	42	0	125
033	06	04	06	58	31	19	29	33	42	52	1	118
034	09	06	05	57	46	17	23	32	20	31	0	112
035	04	05	05	35	29	23	23	28	39	30	1	101
036	06	07	04	37	39	15	32	25	32	36	0	102
037	05	04	08	40	39	20	26	30	29	06	0	099
038	08	09	08	48	44	18	38	40	34	36	0	102
039	08	06	07	48	38	21	30	38	32	33	0	099
040	07	07	08	47	40	14	28	40	31	37	0	119

1	2	3	4	5	6	7	8	9	10	11	12	13
041	09	08	09	31	43	10	25	33	25	30	0	101
042	07	05	09	52	40	18	31	30	31	29	0	112
043	10	09	08	52	54	22	32	39	42	30	0	109
044	08	08	09	34	35	18	29	27	27	31	0	096
045	07	08	09	40	55	21	30	31	35	40	1	118
046	09	09	07	39	44	14	30	37	23	22	0	109
047	09	13	11	50	46	19	40	38	39	49	0	129
048	12	12	10	54	56	17	38	39	39	43	0	121
049	07	04	08	60	47	21	31	43	48	53	1	131
050	11	10	09	42	51	22	30	44	40	46	0	126
051	11	13	08	46	48	18	38	48	41	41	0	129
052	06	11	05	65	52	22	31	45	52	50	1	138
053	10	09	11	44	38	21	40	42	41	42	1	122
054	08	10	08	44	48	20	27	28	37	40	1	116
055	07	09	11	58	53	21	36	45	51	42	1	134
056	10	07	12	61	49	23	36	39	52	52	1	149
057	08	07	08	57	47	21	42	36	44	40	0	124
058	10	09	07	52	56	22	39	48	43	44	1	137
059	10	09	05	49	51	25	39	43	46	56	1	120
060	10	10	10	40	54	22	43	42	36	42	1	115
061	10	09	10	53	62	19	36	43	33	34	0	124
062	10	11	09	64	57	22	34	43	41	43	1	136
063	10	12	12	59	57	24	33	45	45	46	0	121
064	11	12	11	46	60	16	35	42	39	33	1	126
065	07	09	12	54	56	20	39	42	42	41	0	126
066	09	08	10	70	48	24	47	43	43	43	0	134
067	11	06	08	55	48	22	40	32	34	30	1	116
068	09	06	09	52	50	24	34	37	45	48	0	147
069	06	04	11	59	51	25	31	41	49	53	0	132
070	07	07	09	50	52	22	37	43	48	42	0	122
071	08	11	09	63	50	22	39	45	53	58	1	134
072	14	13	14	71	64	26	50	46	52	55	1	137
073	09	10	10	55	44	20	34	47	40	55	1	131
074	16	14	14	72	58	24	52	55	53	59	1	150
075	13	13	16	58	54	25	40	51	52	57	1	130
076	11	09	10	57	25	15	42	35	45	41	1	130
077	09	12	08	56	54	16	46	29	32	40	0	116
078	14	14	09	71	59	24	44	46	54	57	0	140
079	11	12	09	51	51	19	31	35	44	34	0	121
080	10	11	10	59	46	19	33	36	42	51	1	134

1	2	3	4	5	6	7	8	9	10	11	12	13
081	12	10	09	53	47	17	32	44	51	46	0	132
082	11	09	11	58	57	23	39	42	53	52	0	142
083	14	12	13	73	55	19	47	36	52	53	0	132
084	08	11	10	70	51	27	42	44	49	58	0	141
085	10	12	09	72	53	23	39	38	46	55	0	139
086	12	11	11	59	57	23	39	45	38	56	1	131
087	12	14	12	64	58	23	39	49	50	54	0	133
088	11	10	11	65	47	21	29	44	46	54	0	130
089	13	12	14	78	58	25	47	45	57	57	0	150
090	10	13	09	51	49	23	44	47	43	55	0	128
091	11	14	12	37	57	15	32	44	37	24	1	112
092	15	13	12	69	56	21	36	41	49	48	0	119
093	09	09	09	60	59	20	46	45	39	47	0	129
094	10	14	12	65	54	23	42	50	55	46	0	145
095	14	14	11	44	42	16	31	27	38	44	0	111
096	12	11	07	55	63	22	38	50	39	45	0	123
097	12	14	10	54	55	26	43	44	48	44	0	132
098	11	12	12	61	57	22	37	43	43	55	0	117
099	08	08	08	38	50	18	36	39	27	31	0	118
100	08	06	09	40	46	21	28	28	35	36	0	120
101	07	09	10	53	48	24	30	40	31	53	1	118
102	07	08	08	30	47	24	41	48	37	33	0	119
103	06	11	09	52	59	24	35	41	45	53	1	134
104	09	09	10	66	48	24	28	39	40	30	0	110
105	09	08	10	49	45	23	38	45	36	50	0	123
106	06	10	09	35	48	19	28	37	33	40	0	111
107	06	09	08	47	33	18	16	34	28	34	1	117
108	06	05	06	35	37	14	22	43	36	33	1	105
109	09	09	09	54	44	13	36	32	39	40	1	130
110	06	07	05	44	31	22	29	36	39	37	1	131
111	10	12	09	52	30	21	31	35	35	36	1	121
112	10	10	11	64	44	19	43	45	26	36	0	107
113	10	09	12	39	53	20	34	31	29	37	0	119
114	05	05	08	41	43	13	27	31	28	40	1	118
115	08	09	04	39	45	19	28	30	33	27	0	103
116	11	06	08	36	52	22	35	35	38	34	0	110
117	08	07	08	42	45	17	32	25	29	39	1	115
118	08	06	06	33	53	20	24	43	32	42	0	114
119	07	10	09	35	36	14	28	37	34	29	1	110
120	08	07	08	47	50	21	48	44	32	39	0	108

1	2	3	4	5	6	7	8	9	10	11	12	13
121	08	08	09	34	41	15	28	28	39	31	1	120
122	08	07	06	27	42	20	34	34	29	36	1	119
123	11	10	12	71	49	22	46	50	47	48	1	143
124	14	11	11	74	62	24	48	49	50	58	1	147
125	11	10	12	59	54	21	43	42	46	54	0	147
126	13	14	10	71	59	27	45	47	50	56	1	145
127	09	11	10	46	49	24	44	49	51	54	1	147
128	12	13	13	59	59	22	46	49	39	48	1	123
129	12	10	12	62	53	24	44	46	47	50	0	134
130	12	13	12	52	53	24	43	50	49	48	1	140
131	11	10	11	64	54	23	40	43	46	52	1	136
132	12	11	08	59	54	19	44	45	48	52	0	133
133	14	12	12	58	56	23	48	46	48	59	1	150
134	15	13	12	71	58	24	49	50	47	56	0	144
135	12	15	11	70	58	23	26	47	56	57	1	149
136	14	15	13	63	54	18	46	46	51	57	0	142
137	07	08	11	57	47	17	29	38	50	52	0	132
138	14	13	15	67	55	24	51	51	56	55	0	147
139	09	10	11	73	59	13	39	48	48	51	0	135
140	11	11	11	76	58	24	49	42	56	59	1	142
141	12	13	14	83	53	25	43	47	57	60	0	150
142	11	10	12	81	50	24	42	46	54	58	0	147
143	11	12	10	66	55	24	44	43	47	53	0	139
144	12	07	10	50	48	23	42	44	38	34	0	122
145	13	13	12	62	48	20	45	46	48	46	0	140
146	09	10	09	63	47	19	35	33	50	56	0	139
147	10	08	12	57	54	19	47	42	47	50	0	131
148	12	14	11	63	58	22	38	46	46	46	0	131
149	13	11	13	74	58	24	42	40	53	59	0	140
150	14	09	14	63	55	19	43	39	45	44	0	132
151	11	14	12	60	57	23	27	44	45	46	0	129