IOWA STATE UNIVERSITY
Digital Repository

1981

Evaluating an instrument for testout in a business communication course

Carol S. David
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Part of the Educational Assessment, Evaluation, and Research Commons

INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help you understand markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.

2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame. If copyrighted materials were deleted you will find a target note listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in "sectioning" the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.

8122506

DAVID, CAROL S.

EVALUATING AN INSTRUMENT FOR TESTOUT IN A BUSINESS
COMMUNICATION COURSE

Iowa State University   ·   PH.D.   1981

University
Microfilms
International   300 N. Zeeb Road, Ann Arbor, MI 48106

PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy.
Problems encountered with this document have been identified here with a check mark __√__ .

1. Glossy photographs or pages _____

2. Colored illustrations, paper or print _____

3. Photographs with dark background _____

4. Illustrations are poor copy _____

5. Pages with black marks, not original copy _____

6. Print shows through as there is text on both sides of page _____

7. Indistinct, broken or small print on several pages __✓__

8. Print exceeds margin requirements _____

9. Tightly bound copy with print lost in spine _____

10. Computer printout pages with indistinct print _____

11. Page(s) _____ lacking when material received, and not available from school or author.

12. Page(s) _____ seem to be missing in numbering only as text follows.

13. Two pages numbered _____ . Text follows.

14. Curling and wrinkled pages _____

15. Other _____

Evaluating an instrument for testout in a

business communication course

by

Carol S. David

A Dissertation Submitted to the

Graduate Faculty in Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Department: Professional Studies in Education
Major: Education (Adult and Extension
Education)

Approved:

In Charge of Major Work

For the Major Department

For the Graduate College

Iowa State University
Ames, Iowa

1981

TABLE OF CONTENTS

LIST OF TABLES

v

LIST OF ILLUSTRATIONS

CHAPTER I

INTRODUCTION

Background of the Study

The growth of business communication courses

A new kind of college composition course is experiencing rapid

growth as a result of students' desire for practical training to meet

narrowing job opportunities. Institutions offering courses in business

communication have tripled in the past ten years. In 1970, the colleges

and universities with membership in the American Business Communication

Association numbered 150 (Steele 1980); the March 1980 issue of the ABCA

Bulletin reports 501 institutions belonging to the association.

Testout and placement of business communication students

Rapid development of the course has led departments to look for

testout procedures so that students capable of meeting the required

standards can receive credit without taking the course. Although a

written essay is the most direct measure of writing ability (Mehrens

and Lehmann 1978; Purves et al. 1975), time constraints at the beginning

of a term often require departments to screen candidates by objective

tests. Traditionally, colleges have used standardized tests for testout

and placement for freshman English, often in combination with high

school rank. But upper level testout and placement cannot be based on

scores obtained in high school or at freshman orientation. American

College Testing (ACT) reports increasingly lower correlations with

test scores and course grades beyond the freshman year (ACT 1973).

Furthermore, large standardized tests have been faulted for not representing specific populations or specific course goals (Purves et al. 1975). For example, the content of the English subtests of ACT and the Scholastic Aptitude Test (SAT) is designed to reflect the objectives of freshman courses and not those of business communication courses. Thus, interest recently has focused on individually designed objective tests (Johnson 1976), which can be made specific to particular student populations and course objectives. Yet even these objective tests have been criticized in that they tend to measure editing rather than writing skills (Braddock et al. 1963). Furthermore, they require knowledge of standard English (Purves et al. 1975). Finally, their high correlation with academic aptitude shows that they measure in part students' propensity to academic success, which is related to environment. Therefore, they may discriminate against cultural groups (Mehrens and Lehmann 1978; Purves et al. 1975).

Assumptions

1. Qualified students should be allowed to attempt a testout of the business communication course.

2. An objective test designed for the course and student population and properly validated can be a fair and efficient screening instrument.

Problem Statement

In 1976, swelling enrollments in the business communication course at Iowa State University prompted the development of a testout procedure which includes two phases. The instrument that was developed for the

first phase, screening qualified students, is an objective test. Those students attaining acceptable cutoff scores may then move to the second phase of the testout procedure, which is a written essay. If the test is to continue to be used, it is necessary that it be a valid prediction instrument which accurately represents the course and students. The cutoff score which students must attain should minimize placement errors.

This study was initiated to investigate the following questions about the screening process. Does the test predict writing ability adequately? Does the test minimize errors of two types: 1) excluding students from phase two of the procedure who write well but do not score well on the objective test, 2) including students in phase two of the procedure who score well on the objective test but cannot write well? Does the content of the test represent the course?

Objectives

1. To examine the correlation of the business communication objective test with the criterion measures of essay grade and course grade.

2. To examine the data for the purpose of establishing a cutoff score that will minimize two types of errors of placement: 1) the exclusion of students from phase two of the testout procedure (writing an essay) who can write well but cannot score well on an objective test of mechanics and style, 2) the inclusion of students in phase two of the testout procedure who score well on the objective test but cannot write well.

3. To examine if the content of the objective test fairly represents the course and students.

Special Terms Used in this Study

The term "grammar," for this study, will denote traditional grammar, the system with which most students are familiar, and not the newer structural or transformational systems. Traditional grammar is used here to denote a system of rules used to prescribe correctness in sentence structure and punctuation. In contrast to "formal grammar," which signifies the systematic description of an entire grammar system, the grammar taught in business communication classes is confined to the correction of errors on papers and brief discussions in class.

The terms "grammar" and "mechanics" are used synonymously, except where "grammar" labels a subtest which includes items on agreement and sentence structure. In the conventional terminology of English testing, mechanics items on objective tests are usually designated "usage."

For this study, a teaching method in courses which focus on grammar and include no actual writing will be labeled the "direct" teaching method.

Limitations of the Study

This study was conducted within the framework of the normal business communication college course. No controls were placed on teaching or grading procedures. Because the criteria for course grades were set by individual classroom teachers, wide variations in grading may occur. The result may be a diminishing of the correlations made with the objective test and course grade.

Design of the Study

A one hundred-question objective test primarily of grammar, prepared by the staff of business communication at Iowa State University, has served as an instrument to screen students for testout. The validity of this instrument was tested by correlating it with two criteria: 1) writing sample, 2) course grade. These correlations must be judged adequate if the test is to continue to be used. To validate the test, this study used thirteen sections of business communication students who received the test at the beginning and end of the course the spring of 1979. Teachers submitted final grades, and these and four sections of final essay grades were correlated with the test. The following fall the test was again administered to three classes to provide a replication of the validation process. At this time, essays were collected at the beginning of the course as well as at the end. This replication provided correlations between test, corresponding essays, and final grades.

For the content areas of the test to be judged useful, each subtest should correlate adequately with the criterion measures and be representative of the course objectives. In examining the content of the test, this study correlated the subtests with the criteria of writing sample and final grade. Furthermore, the test was administered as both pretest and posttest in order to observe the gains. Principles showing no gains may not be relevant to the course and could be dropped. But failures to show gain may be caused by the principles being taught poorly or the concepts being too difficult to learn. Because grammar is not taught directly in business communication, it is

difficult to judge if the principles represented by subtests are relevant to the course. Therefore, as a check on gains when grammar is taught directly, a group of students in business writing short courses was administered a shortened form of the objective test before and after the course. A further analysis of the teaching and learning process was available with this group.

Subjects

Business communication students   College juniors and seniors enrolled in thirteen sections of business communication at Iowa State University comprised the subjects in the business communication group. These students may have no special interest in or talent for writing, and they may have had little opportunity to write during college. Both characteristics make them representative of the general young adult population and not of upper level creative writing students.

Short course students   Sixty Iowa State University employees voluntarily enrolled in short courses given by the business communication staff constituted the group which received direct grammar instruction. These students have educational backgrounds similar to the college students except for their additional experience in the working world. They are secretaries who do not write but edit others' work and professional and nonprofessional staff who compose their own correspondence.

Courses

Business communication course   Business communication at Iowa State is a one-quarter course giving instruction in the typical writing

of business settings. Unlike freshman English courses, which include wide varieties of writing experience, from personal narrative to term papers, all of the practice for this course is in writing which James Britton terms "transactional" (Britton 1975), communication between the writer and a known audience for a given purpose. Basic textbooks offer sections on content, organization, tone, style, and mechanics (Wilkinson, Clarke, and Wilkinson 1980; Lesikar 1980; Wells 1977; Treece 1978), which set the criteria for evaluation of this writing. Consequently, the grading is somewhat more uniform than in freshman English, where staff members have more freedom to emphasize areas they deem important.

Adult short course   The six-hour short course, half of which is devoted to teaching grammar, is presented by four teachers from the college business communication staff. Adults receive instruction in composing but are given no writing experience. The course is primarily devoted to teaching grammar. It is held on consecutive days, offering three hours of instruction each session.

The instruments

The full objective test   Since 1976, a one hundred-question objective test prepared by the staff of business communication at Iowa State has been used as a screening instrument for testout for the business communication course. Those students receiving scores of 85 or above are then asked to write two essays, which are evaluated by two teachers of the course. The test consists of eighty-five true-false questions on spelling, punctuation, and grammar, and fifteen multiple choice questions on style and tone.

<u>The shortened test</u>   A shortened form of the test, comprised of twenty-eight items taken from the full test, was used in this study to measure if students improve their score after instruction. One short course group was given the same test before and after the course, and another group received an alternate form of the test as a posttest.

<u>Questionnaires</u>   A questionnaire was administered which sought college students' permission to obtain ACT composite and English scores and to gain information on students' family educational background. In addition, students marked Likert-type scales designed to measure writing confidence and motivation. Short course students marked similar scales on confidence.

Methods of data collection

Nine teachers of business communication administered the objective test to their students during the first and final weeks of the course in spring 1979. A questionnaire was distributed the third week of classes asking students for permission to obtain their ACT scores. Those students who gave permission to use ACT scores in the study were recorded and their scores obtained from the university records office. Essays were collected from four sections and graded by three teachers of the course. The following fall the test was administered to three sections to replicate the study. Essays were collected from the students at the beginning and end of the course and graded by two teachers.

Four teachers presenting the short courses to university personnel cooperated in administering the shortened form of the objective test at the beginning and end of each course. One group was tested again

one month later by delivering the test to each individual and collecting
it the following day.

Analysis of Data

The Pearson product-moment correlation between objective test and
the criterion measures of essay and course grade was the statistical
method used in computing objective one. Objective two, setting cutoff
scores and determining the placement error, was achieved by examining
scatterplots to obtain the figures representing the percentage of stu-
dents misplaced by a given test score. Objective three, analyzing the
content of the test, was accomplished by correlating the subtests with
essay and course grade. Furthermore, intercorrelations of subtests and
full test provided evidence of the homogeneity of test areas. Another
check on the content compared pretest and posttest scores to identify
areas showing gain. Comparisons of gain scores for college students and
short course students were made by using a group t-test.

Organization of the Paper

Chapter II, the review of literature, cites pertinent literature
from placement testing and testing in composition in order to suggest
methods of validating the test and to present correlations found in other
research. Evidence both for and against the use of an objective test
as a measure of writing is included in this review. Chapter III con-
tains a complete explanation of the subjects and instruments used in
the study as well as the essay grading procedures for both the main study
and the replication conducted the following year. Chapter IV, the

findings, contains the results of the computer analysis of the data.

Chapter V is a discussion of the findings and presentation of the con-

clusions.

CHAPTER II

REVIEW OF LITERATURE

The purpose of this study was to evaluate the business communication objective test as a screening instrument for testout procedures in the course. A review of the literature of both prediction tests and testing in English composition identifies capabilities and limitations of using objective tests for placement in composition classes. This chapter includes methods of evaluating placement tests and pertinent scores and statistics which earlier studies have found. The first section discusses requirements and methods for the validation of prediction tests. The second section discusses cutoff scores and the resulting placement errors. The third section discusses the content of prediction tests.

Prediction

Background

A screening instrument used in testout procedures for a composition course must discriminate those students who are capable of writing at a level comparable to students who successfully complete the course (Pearson 1978). The test is a "prediction" test in that it predicts the writing capabilities of students so that they may be divided into two groups: those who have likely achieved the objectives of the course and those who likely have not. In the business communication testout procedure, those who are identified as likely to have the necessary skills will then write an essay in which they can demonstrate their ability to

achieve the aims of the course.

The essay represents the achievement test in the placement procedure. The objective test represents the predictor of writing capability and is used prior to the essay test because of the impracticality of administering to and grading essay tests for large student populations. Particularly at the beginning of a term, essay tests require teachers to undertake a heavy grading load. Thus, the machine-graded objective test serves to narrow the population taking the essay test.

The effectiveness of such easily evaluated tests as substitute measures of writing ability lies in their correspondence with the primary task of writing. Here "correspondence" of two tasks designates the extent to which levels of performance in both tasks are related. In evaluating the correspondence, it is necessary to first choose measurable tasks representing the abstract quality, writing ability. These tasks are called criterion measures (Thorndike and Hagen 1977). Appropriate criterion measures for writing ability could be the following: 1) graded essay, 2) course grade in a composition class (Mehrens and Lehmann 1978; Nunnally 1970). If students' scores on the test correspond to essay grades or course grades, whether both are high or low, the test can substitute as a measure of writing ability.

The technical concepts of reliability and validity enter the quantitative study of such tests and their association with the criterion measures. For a better understanding of these terms, which will be used in evaluating the objective test in this study, discussions of each are given below.

Reliability

Reliability of a measure is the degree to which levels of perform-ance on that measure are repeatable (Thorndike and Hagen 1977). Relia-bilities of computer-scored tests are easily calculated. Kuder-Richardson Formula 20 is commonly used and accepted as a conservative estimate (Thorndike and Hagen 1977). The formula is an estimate of the correlation among items on the test and between items on the test and a related hypothetical test. In general, the higher the reliability of the test, the better. Standardized tests obtain reliabilities of .80 to .90 (Nunnally 1970). Reliability is related to measurement error, which is the discrepancy between an individual's true score--the score achieved if no measurement error existed--and the individual's obtained score on the test (Nunnally 1970). For prediction tests, the importance of the test's reliability lies largely in its effect on the correlation between the test and the criterion measures. Lower relia-bility of either test or criterion will depress the correlation because scores will fall randomly on the plot of the two measures and blur the actual correspondence of the two. Thus, adequate reliability of both test and criterion is desirable for obtaining adequate correlations (Mehrens and Lehmann 1978), but the correlation is the crucial statistic (Nunnally 1970).

Validity

A test's validity represents the degree to which the test measures what it is intended to measure (Thorndike and Hagen 1977). For example, the screening instrument for the business communication testout must

predict business writing ability. The predictive validity is the most crucial concern of a placement test (Nunnally 1970). The predictive validity of the objective test as a measure of writing ability is quantified by the correlation coefficient obtained when the test scores are correlated with scores on the criterion measures. This correlation is made by having students perform both tasks and observing whether the students tend to receive similar numerical scores on both tests. The Pearson product-moment correlation coefficient r represents the most commonly used method of assessing the extent to which this happens (Mehrens and Lehmann 1978). The correlation has values near one if the levels of performance on the two tests are closely related and values near zero if they are not. If the relationship between the two types of tests is high (r near one), the objective test is judged a good surrogate for the essay test or course grade. If the relationship is low (r near zero), the objective test cannot serve. The adequacy of the correlation coefficient is judged by comparisons with coefficients found for other similar tests and their criterion measures (Thorndike and Hagen 1977).

In addition to predictive validity, content validity is an important concern of a prediction test (Purves et al. 1975). The content of a test may be measured qualitatively (Nunnally 1970). On a test of academic capabilities, the content should be representative of the course objectives (Mehrens and Lehmann 1978). However, items on prediction tests, unlike achievement tests, need not proportionately represent the course content (Nunnally 1970). They must test skills required for achievement in the course. While a skill such as spelling may not be taught in the

composition course, it may predict success in the course and therefore be a suitable area for a prediction test to include. However, test designers must ensure that such skills do not discriminate against individual students or groups of students who have not had access to obtaining them. To avoid such discrimination, test designers should base the test as closely as possible on the objectives of the course and the characteristics of the student population being tested (Johnson 1976). The best judges of the appropriateness of the test content are teachers familiar with both the course and students (Nunnally 1970).

A first step in evaluating a test is to obtain reliabilities of the test and the criterion measures of writing ability, graded essays and course grades.

Obtaining Reliability

Variables

High reliability is achieved by reducing the random variability that may occur in the four identifiable components of a testing situation: 1) the test, 2) the testing environment, 3) the student being tested, and 4) the test scoring (Nunnally 1970). Both an objective test and its criterion measures are open to variation in these four areas.

The test   A test will admit random differences in performance if the items on the test are not representative samples of the content area of the course. The more the items reflect the essential content of the course, the less chance enters the students' abilities to answer correctly. Including a sufficient number of items on the test will in-

crease the likelihood of a fair sampling of the course content (Nunnally 1970). Essay tests, which offer only a single topic, are especially subject to variation in students' performance. To avoid giving some students unfair advantage, testers should ensure that essay topics are familiar to all students writing the test (Braddock et al. 1963).

If a test allows opportunities for guessing, a lower reliability may occur. Open response answers would decrease chances of guessing, but such tests are more difficult to score (Nunnally 1970). On the other hand, a test format with true-false items allows the greatest opportunity for guessing. However, such a format may contain more items, because true-false items require less of the student's time to complete (Mehrens and Lehmann 1978). Increasing the number of items will usually improve the reliability of the test and counterbalance the error induced by guessing (Nunnally 1970).

The test environment   Nonuniformity in the test environment will allow for changes between individuals' performances. Time limits and instructions should be standardized so that each time the test is administered, conditions will be identical (Nunnally 1970). The degree of students' comfort and familiarity with the surroundings should be constant.

The student being tested   Changes in individual performances, which test administrators cannot control, will affect a test's reliability. A student may have a bad day or may be distracted by outside problems. The more the test score is influenced by such changes, the less reliable the test. Writing ability may be particularly susceptible to mood, and therefore two essays written at different times are a better

measure than one essay (Diederich 1974).

Test scoring   Essay grades and course grades are influenced by subjectivity in the grading process to which machine scored tests are not liable (Nunnally 1970). Teachers do not always agree on what constitutes good writing, so their essay grades may not be based on similar criteria. Course grades may show even greater variability because some teachers include factors such as class participation and objective tests in the final grade. Teachers may not be consistent in their own grading. They may show more leniency on one day or to one class.

Achieving reliability in essay grading has been given much attention in the literature in the past twenty years. Recent research points to the importance of raters agreeing upon their rating criteria in order to achieve adequate reliability of rating (Cooper and Odell 1977).

Attendance to controlling these variables will aid testers in achieving acceptable reliabilities. Most of the requirements are readily attainable in placement testing. The exception may be the allowance for writing essays on separate days.

Methods of Essay Grading

Paul Diederich, senior research associate of the Educational Testing Service (ETS), reported on a 1961 study testing reliabilities of ratings obtained without raters formulating their criteria (Diederich 1974). His group asked sixty judges chosen from six different fields to grade 300 student essays and rate them from one to nine. One hundred one essays received every rank; no essay received less than five dif-

ferent rankings. The correlation between raters was .31.

A recent grading study conducted by Wilkinson (1 ,9) asked teachers of business communication to rank seven letters. It then compared the rankings to those made by a seasoned teaching staff at the University of Alabama. Although the raters agreed on the two worst letters, the three letters considered best by the Alabama staff were ranked 4.5, 3, and 4.5 (out of 7) by the outside raters. The outside raters also ranked first the letter receiving a rank of 5 from the Alabama staff. Ratings assigned by outside raters showed about twice the variability of ratings by Alabama teachers. The Alabama teachers' ratings were more consistent because the teachers were accustomed to working together and consequently focused on similar standards.

To achieve reliability of essay grading, raters should be teachers from a single department who are accustomed to working together. These teachers should then set up the criteria for rating (Cooper and Odell 1977; Diederich 1974). Two raters are adequate for obtaining reliabilities (Mehrens and Lehmann 1978), and the grades from each may be correlated using the Pearson product-moment correlation coefficient. A coefficient of .50 is typical (Diederich 1974) for a correlation of two raters. This figure may be corrected by the Spearman-Brown prophecy formula, boosting an average of .50 to .67. Reliabilities for essay grading in this study were calculated this way and coefficients of at least .67 are required for essay grades to be considered satisfactory criterion measures.

The analytic method of rating, where raters give thorough readings to each paper and often mark comments or check grids, offers the most

accurate ratings, but holistic grading is capable of nearly the same accuracy and is more efficient (Cast 1939). In holistic grading, raters read quickly through the paper and award a grade on the basis of a general impression of how well the paper meets the criteria initially set. Raters make no markings and spend no more than two minutes per paper (Braddock et al. 1963). Raters who have agreed on determining grading criteria have achieved reliabilities as high as .90 using holistic grading (Cooper and Odell 1977). This method has become widely accepted (Cooper and Odell 1977), and this study will implement it.

Correlation of Objective Test and Essay Grade

Several studies which used holistic grading have found good correlations between essay grades and objective test scores. The College Entrance Examination Board (ETS 1979) reports correlations of .48 and .50 between the SAT English composition test and graded essays. Some studies have found higher correlations. Nisbet (1955) found objective test and written essays graded holistically to correlate .82. With such high correlation, Nisbet concluded that the objective test alone was sufficient for prediction. An essay need not be added to the procedure. Similarly, Noyes (1963) reported on a special study for the College Entrance Examination Board (CEEB) on the SAT English composition test. He stated that the test correlated so highly with essays that an essay added to the objective test would raise the predictive ability negligibly.

On a department-designed test developed for placement at Northern Virginia Community College, Johnson (1976) found correlations between test and essay to be .62-.65. She concluded the objective test was

sufficient for classification without requiring an essay. A replication

of her study showed lower correlations of .42-.46.

However, some critics object to testing by objective tests alone

even when such correlations are high (Purves et al. 1975; Braddock et al.

1963; Weaver 1979). The objections are of two types. One is concerned

with the fact that objective tests do not test actual writing. Instead,

they require the recognition of errors in others' work. The result of

such testing may lead both students and teachers to overemphasize the

importance of grammatical errors in the writing process and underempha-

size other areas of composing. A second objection is directed at using

large standardized tests for placement. The association between the

test and the writing ability of a specific group may not be close enough

so that the test can serve as a fair predictor for that group. Depart-

ment-designed tests such as the business communication test in this study

are not open to this second criticism when they are based on course

objectives and errors typical of the student population (Johnson 1976).

Correlation of Objective Test and Course Grade

Like written essay, grade received in a composition course is a

measure of writing ability. One advantage of course grade is the

opportunity it allows students for writing on a number of topics at

different times. This variety reduces the chance of students receiving

lower essay grades because of unfavorable topics. But a limitation is

the rater variability. A personal relationship between teacher and

student may influence the grade, and this element is virtually im-

measurable (Diederich 1974). Furthermore, criteria for grades may

represent more than writing ability, by including tests and class performance. Noyes (1963) reported that SAT avoided correlations with grade in validating their English composition test because of high variability of grading criteria. Course grade as a measure of writing ability offers less reliability than essay grade and may result in a lower correlation coefficient. Both ACT and SAT English composition tests correlate about .40 with course grades in freshman English (ACT 1973). However, a common association between course grade and objective test influences the correlation and obfuscates the results.

Bias of objective test and course grade correlation

Grades and scores on tests measuring academic abilities are influenced by the cultural climate in which students have grown up (Thorndike and Hagen 1977; Mehrens and Lehmann 1978). Grades are based on students' interest and effort as well as students' rapport with the teacher (Thorndike and Hagen 1977). These qualities and abilities are more often nourished in a home of educated parents; thus, parental education becomes an important factor in predicting academic achievement (Brown 1978). In the same way, performance on tests such as those in English composition may be related to similar environmental factors. Aptitude tests are not pure measures of intellectual potential but measures of abilities developed by the same factors which develop academic achievement in students (Thorndike and Hagen 1977). These common associations may result in unrealistically high correlations between test and grade. Thorndike and Hagen (1977) give correlations of above .50 as typical.

The results of several studies of placement tests in the 1930s

may illustrate this association. Because the concept of cultural bias was not popularized at the time, these studies gave it no mention in their findings, although the associations of test score and grade suggest that the bias obtains. In the 1930s, before standardized tests dominated placement in English, departments often designed their own tests and found that students scoring high on the objective test were apt to achieve high course grade. The following studies illustrate this association.

In the first of these studies, a placement test at Purdue University, McKee (1934) found that the freshman students scoring high on an objective test and low on an essay tended to attain higher composition grades than those scoring high on the essay and low on the objective test. Correlations were not given, but the conclusion was that the test identified high achieving students especially well.

In the second of these studies, a test battery taken by 1,500 college freshmen at the University of Chicago, Stalnaker (1933) found similar results. He compared scores on an objective qualifying test with an essay test and investigated cases where wide discrepancies in scores occurred (students who scored above the 75th percentile rank on one measure and below the 25th percentile rank on the other). He found, first of all, that only thirty-eight cases of wide discrepancy occurred. He then asked teachers how well the students had performed in composition classes. He found that those students with high objective tests scores and low essay scores performed better in classes than those with high essay, low objective test scores. Stalnaker also noted that students with high test-low essay scores tended to achieve higher on intelli-

gence tests which had been administered. Scores from the objective test
correlated .69 with scores on the intelligence test, indicating a high
association between scores on aptitude tests of English composition and
general academic aptitude.

Figures from both of these studies suggest that because high objec-
tive test scores are associated with good class performance and general
scholastic aptitude, objective tests measure propensity to academic
success as well as writing ability. A third study which gives correla-
tions of test and grade offers further evidence of this bias. At Iowa
State College, Starbuck (1933) studied the effects of a placement test
for freshman English which his group designed on the basis of errors
observed in freshman themes. The department classified students using
a combination of test, writing sample, and general aptitude scores, and
correlations for each with final grade were given as follows: objective
test .49; essay .36; aptitude .52. The aptitude test correlated highest
with course grade. The objective test correlated higher with grade than
the essay. The lower associations of essay and grade reported by this
study and the McKee study may have resulted from the unreliability of
essay grading at this early date. No mention is made in either of the
studies of formulating criteria for grading, so it might be expected
that low reliabilities which were obtained were caused by variations
in criteria.

The aptitude test used in the Starbuck study correlated .58 with
objective test and .56 with essay grade. The three tests together,
objective, essay, and aptitude, correlated .61 with final grade. The
results show that academic aptitude not only correlates highly with

grade but also, when added to essay and objective test scores, boosts the correlations of those scores and final grade substantially.

ACT composite test score, which is defined as a measurement of developed academic abilities, correlates .72 with general intelligence test scores, and the English composition test correlates .60 with intelligence test scores (ACT 1973), further evidence of the close association of all aptitude tests.

Because of the high association of objective tests and general academic ability, course grade should not be used solely as a criterion measure of writing ability.

This study correlated the objective test with course grade as a second check on the predictive ability of the test. It attempted to identify the degree of cultural bias in the test by first correlating test scores with scores on the ACT composite test, which reflects general academic aptitude. It then examined associations of test scores with students' level of family educational background for evidence of cultural bias.

While no test is completely free of bias (Mehrens and Lehmann 1978), efforts can be made to reduce the bias by fitting the test closely to the content of the course and to the student population (Johnson 1976).

In summary, students who show high competency of performance on objective tests are usually students who also show general academic abilities. Both abilities have the common association of family educational background, which fosters such success. For this reason, the correlation between test and course grade may be unrealistically high.

The Use of the Test

If a placement test has been found to have acceptable correlations with the chosen criterion measures, the test's predictive validity may be assumed satisfactory. However, the fairness of the test may be judged by examining how it is used (Mehrens and Lehmann 1978). For instance, screening instruments may be used alone or in conjunction with other measures, such as high school rank. Of primary importance in judging the use of the test is the score set for students to attain. Students not attaining this cutoff score are excluded from college admissions or excluded from upper level courses. In the case of the business communication testout procedure, students not attaining the cutoff score are not allowed to write the essay and dropped from the testout procedure. Some placement error will exist in the use of any test (Mehrens and Lehmann 1978), and the inequities must be reluctantly accepted.

Prediction error

For the use of the test to be judged fair, procedures must be designed to minimize the placement error. One way to achieve more accurate prediction is to use several instruments. A battery of tests is better than one (Nunnally 1970). But when departments must use just one prediction instrument, they are faced with finding a fair cutoff score. Two types of placement errors are possible: 1) excluding students who could find success, 2) including students who cannot succeed (Mehrens and Lehmann 1978). The cutoff score must be set low enough to include students who are unable to score well on objective tests but who have potential for success in writing. Yet it must

not be so low as to include students without potential for success in writing. The error of the first type is unfair to students. The error of the second type is time-inefficient for departments. Excluding students who write well but cannot score high on an objective test, the first type of error, has been recognized in the literature of placement testing, but it is difficult to quantify. These are students who do not meet the cutoff scores and are excluded from the placement procedure (Mehrens and Lehmann 1978), thus remaining uncounted. Quantitative data on this group are unavailable in the literature. McKee (1934) recognized the group qualitatively by acknowledging that his objective placement test missed some students with writing ability who are good at the skills the test does not measure and not good at skills the test does measure.

On the other hand, students who pass the screening but fail to find success in writing are more easily identified. Tables in a study at Iowa State University on a freshman testout procedure (Pearson 1978) gave percentages of students meeting the screening scores but failing to pass the second phase of the procedure, the essay test. The Minnesota Scholastic Aptitude Test (MSAT) was used as the prediction instrument because it was found to be a better predictor than either ACT or course grade. Figures showing the percentages of students failing to pass the essay test at given levels of achievement are shown in Table 1. Of high-achieving students with scores of 50 to 58 (B range), which was chosen as the cutoff score, only twenty percent met the testout requirements.

Table 1. Placement error for testout in fresh-
man English using MSAT scores[a]

Score category	Testout result	
	pass	fail
A (59-up)	.39	.61
B (50-58)	.20	.80
C (37-49)	.09	.91
D (29-36)	.00	1.00
E (28-below)	.00	1.00

[a]Reproduced by permission of Iowa English Bul-

letin. Copyright 1979 by the University of

Iowa, Iowa City.

The figures in Table 1 clearly indicate that the prediction error shown

by standardized tests is considerable.

In summary, no prediction test will correlate perfectly with the

criterion measures, and some error will occur in the screening process.

Test administrators must not set cutoff scores so high that a large

number of qualified students are excluded from the second phase of the

testout procedure. On the other hand, a low cutoff score will include

unqualified students in the second phase of the procedure and give

teachers the burden of grading an excessive number of essays. Adminis-

trators must choose a score that is fair to students and practical for

departments.

Content

The content of an objective test predicting writing ability must measure skills associated with writing. However, the test must select a limited number of areas associated with writing. A test of English usage has historically served as a predictor of writing ability even though it actually measures editing rather than writing ability (Diederich 1974). "English usage" signifies a test of the conventions of standard written English (CEEB 1978). Most items on English usage tests examine grammar, idiom, and diction, although some tests may include broader areas of style, such as parallelism and clarity, under usage. The SAT Test of Standard Written English contains thirty-five questions on grammatical usage. Another fifteen-item subtest is labeled sentence correction and pertains to principles of grammar and style (CEEB 1978). The seventy-five-item ACT English Usage Test contains some questions on style, logic, and organization, but most items test punctuation, grammar, sentence structure, and diction (ACT 1979).

Usage is not the only area of composition which objective tests of composition can contain. The National Council of Teachers of English report (Purves et al. 1975) urges testers to examine other areas of writing, such as style, content, and organization. Since style is related to the purpose and audience of the writing, department-designed tests are better able to examine style appropriate for the assignments of a specific course. Standardized tests are limited to essay style, which would not be typical of business writing. The business communication objective test contains fifteen items testing style and tone.

Prediction tests need not contain items which proportionately repre-

sent a course. If the skills tested are good predictors of writing, the test may be useful. However, an achievement test which contains items closely associated with course objectives may prove to be a good prediction instrument (Mehrens and Lehmann 1978). The business communication test in this study contains eighty-five questions of English grammatical usage, which are divided into three subtests of spelling, punctuation, and grammar. Grammar is only one of five categories making up the criteria for evaluation of the writing in the course, so grammar receives a disproportionately heavy emphasis on the test. But business correspondence requires correctness in grammar, and errors on assignments are judged severely (Warner 1979). The dominant role of grammar in the test may be justified if grammar proves to be integrated into the course objectives.

One way of determining if the content areas included in the test are representative of the course objectives, and thus fair to include on the test, is to administer the test as a pretest at the beginning of the course and posttest at the end and observe the gains (Nunnally 1970). Areas showing no gains may not be learnable in an academic course or they may not be clearly taught. Grammar is not directly taught in the business communication course. To test if grammar is learnable in the quarter's course, this study used a shortened form of the objective test administered to a short course group where grammar was directly taught. Since questions were chosen for their high level of difficulty, any gains shown from pretest to posttest will indicate the grammatical principles are accessible to students when presented in the classroom.

Gain scores must be interpreted with caution. If the pretest and posttest are highly correlated, students' performances are similar on both. The result is that the diminished variability will reflect in the gain scores and lower gain score reliability (Mehrens and Lehmann 1978).

Another factor affecting the interpretation of gains is that gains do not occur uniformly across all scores. High scorers have limited possibilities for improvement (Borg and Gall 1979). On a test of one hundred items, a score of 96 can move only 4 points higher. The result of this "ceiling effect" is that lower scores have a greater potential to rise than higher scores. The level of difficulty of the test will influence the magnitude of the ceiling effect. An easy test will limit gains for a number of students, while a difficult test could allow room for all students to gain. Good students are further hampered by the diminished pool of knowledge from which they must learn. Poor students, on the other hand, have a large pool of information still remaining, much of which may be relatively easy to learn. Therefore, the achievement of a true gain of 99 to 100 is greater than that of 75 to 76.

The phenomenon of regression toward the mean represents another problem in interpreting gain scores. If a student scores very high or very low on a test, it is likely that the performance is not typical. On another test the score would be less extreme, drawn towards the middle (Borg and Gall 1979).

The above factors are largely directed to the interpretation of individual gains and indicate that students at the same level of achievement are more usefully compared. This study uses group means.

Any change from pretest to posttest will verify only that learning does or does not occur, and for this limited interpretation of group means, gain scores provide adequate accuracy (Mehrens and Lehmann 1978). In comparing group gains, average gain scores will be more meaningful if pretest mean scores of the groups are similar.

However, the use of a pretest will affect gains. The repetition of taking the test is likely to raise the posttest score. Students receive experience in taking the test and are sensitized to the principles involved. Having been exposed to the principles on the test, they may then be aware of what they must look for in the instruction. In this study, college students were not aware that they would be retested. Although they may be alerted to the principles on the test, they would not be motivated to learn them by the possible reward of an improved grade. Similarly, short course students were not aware that they would retake the test. The proximity of the tests administered on consecutive days, however, suggests that students may be sensitized to the test's principles.

Correlation of subtests with writing sample and grade

Correlations between subtests and the criterion measures will show the predictive power of the test's subject areas. Studies have found that items of usage generally correlate well. In a test administered to 3000 students at Oberlin College (Harston 1930), subtests and grades correlated as follows: capitals .291-.373; grammar .473-.486; punctuation .387-.508; sentence structure .331-.443. (The study separated subjects by sex, and the data on women are shown in the first figure.)

The total mechanics test correlated .433-.634 with grade. Reliabilities of test and subtests were not given.

Three other studies correlated subtests with essay grades. In the first study, an objective test was administered to 112 students in grade twelve (Edmiston and Gingerich 1942). Correlations of subtests and essay were as follows: usage .55; capitals and punctuation .23; sentence structure .37; spelling .26. Reliabilities of the test and subtests were not given, and because the essays were graded by just one rater, the reliability of the essay grades is questionable.

In a study of the SAT English composition test (Noyes 1963), where the reliability of essay grading was high, usage correlated .707 with essay grades. Subtests on areas other than mechanics, such as one on logic and tone and one on organization, were lower, .568 and .458 respectively.

Johnson (1976) gave the following correlations of subtests and essay grades for her freshman English placement test: spelling .53; punctuation .58; content .37; organization .24; and style .47. Reliabilities of the business communication objective test and subtests and correlations of subtests with essay and course grade will be given in Chapter IV.

Subtests of usage have shown higher correlations with the criteria of written essay and course grade than subtests measuring other areas of writing. However, to score well on a test of usage, a student must have a knowledge of standard English. This requirement may represent a bias against cultural groups in standardized tests for general writing courses which allow variations in usage and style on writing assignments.

Test bias

Besides their failure to represent specific course objectives, standardized tests are removed from their target group and norms may not fairly represent these groups (Purves et al. 1975). They may test material inaccessible to or inappropriate for a specific population. A department-designed test for which teachers familiar with the course submitted questions will more fairly represent the student population (Nunnally 1970; Johnson 1976). The business communication test was designed in this way. However, tests requiring a knowledge of standard English will discriminate against cultural groups (Purves et al. 1975). In general composition courses, limiting questions to standard English may be unfair since writing assignments may allow for a flexibility in usage. But in business communication no such flexibility is allowed. Adherence to standard English is one of the course objectives (Warner 1979; Wilkinson 1977).

Correlation with ACT English usage test

Another way of checking on the content of a test is to correlate scores with those from another test measuring the same abilities (Nunnally 1970). This study will use ACT English Usage Test scores for this correlation.

Summary

The most important statistic in the evaluation of a screening test of English composition is the coefficient obtained in the correlation between the test and its criterion measures (Nunnally 1970). This co-efficient represents the predictive power of the test.

Because lowered reliability of either test or criterion measures will attenuate the correlation, it is important to reduce the variations in the testing situation in order to raise the reliability of both (Nunnally 1970). Test content, test environment, the student tested, and test scoring are all liable to chance variation. While scoring of objective tests is consistent, scoring of essays is especially open to qualitative judgments. To achieve adequate reliability, raters should agree on the rating criteria before grading the essays (Braddock et al. 1963; Diederich 1974; Cooper and Odell 1978). With this agreement, typical reliabilities have been about .67 (Diederich 1974).

Essays represent the most direct measure of writing ability (Mehrens and Lehmann 1978). If the test scores correlate well with grades on essays written contemporaneously by the same student population, then the test can serve as a surrogate measure of writing ability. Grade in a composition course can also serve as a criterion of writing ability, but because grading lacks uniformity and may include factors beyond writing, such as attendance and other tests, the correlations may be lowered. But grades and aptitude test scores have a common association with students' general academic ability and family background, which may raise the correlation unrealistically. It is necessary to consider this bias in interpreting the correlation of test scores and course grades. Correlation coefficients of .42-.85 have been reported for essays and objective tests (Nisbet 1955; Diederich 1974; Johnson 1976; ETS 1979). Coefficients from .40-.50 have been obtained for course grade and objective test (Starbuck 1933; ACT 1973).

If a screening test proves an adequate predictor, cutoff scores may

be set. Because no test is a perfect predictor, errors of two types will occur: 1) students who do not meet the cutoff score but can write well will be dropped from the testout procedure, 2) students who meet the cutoff score but cannot write well will be included. Indications from the literature are that the second type of error is about 80% (Pearson 1978). No figures are given in the literature for the first type.

An evaluation of an objective test must also examine the content of the test. Test designers have been urged to include subtests on style, organization, and content as well as grammatical principles as a part of objective tests (Purves et al. 1975). However, subtests on grammatical usage have shown high correlation with essay grades. Coefficients of .50 and above have been reported (Edmiston and Gingerich 1942; Noyes 1963; Johnson 1976).

A test of grammatical usage will include a cultural bias because it demands a familiarity with standard English, to which all cultural groups have not had access. While such a requirement may be less fair for courses in general composition, the requirement is consistent with the course objectives of business communication (Wilkinson 1977; Warner 1979).

This study evaluated the business communication objective test using the criteria found in the literature. Methods for the evaluation are explained in Chapter III.

CHAPTER III

PROCEDURES

The business communication objective test was administered at the beginning and end of spring quarter 1979 to all students enrolled in the course. A replication of the study using three classes was carried out the following fall. In addition, pre and post essays were collected from samples of students. In an actual testout process, students not meeting the required test score are not allowed to write an essay and are dropped from the procedure. By including all students and not just the high-scoring students, this study could analyze corresponding test scores and essay grades for all levels of performance, including those with low test score-high essay grade.

The test content is comprised of eighty-five items on grammatical usage; however, grammar is taught only indirectly in the course. To check if the difficult grammatical principles on the test are learnable in an academic course, the study included business communication short courses where grammar is taught directly. Students in these courses received a shortened form of the objective test before and after the two-day course. Thus, the study used subjects in regular college credit courses. A supplementary group of students in noncredit courses was included to check if the content of the test is learnable.

This chapter will discuss the following: 1) subjects, 2) courses, 3) instruments, 4) data collection procedures, and 5) methods of statistical analysis.

Subjects

Business communication students

Juniors and seniors enrolled in business communication courses at
Iowa State University during the spring quarter of 1979 comprised the
subjects of the major study. Of the approximately 300 students enrolled
in thirteen sections, 258 are included in the pretest data. Two hundred
eight students were matched with pretest, posttest, and questionnaire.
The loss was caused mainly by absences at the time the pretest, posttest,
and questionnaire were administered. Each was given on a separate day,
and no makeups were allowed. Ten students dropped the course and were
removed from the study. Two students failed to complete the posttest
and were dropped.

Most students taking this course choose it under a communication
requirement. The largest number of students major in business and
agriculture, but some major in home economics, education, and science
and humanities. Because students are taking the course as a require-
ment for graduation and not as an elective, they may have little
experience in writing. They are, however, career oriented. Nearing
graduation, they are well motivated to learn the practical skills of
business writing.

Short course students

Iowa State University professional and nonprofessional staff who
have elected to participate in short courses in business communication
offered by the university personnel department comprised this group. The
short courses, taught by teachers of the business communication staff,

are given on a continuing basis two or more times a year. In this study,
thirty-three people attended courses held in the late spring and early
summer of 1979. Most were secretaries, clerks, and office personnel.
This sub-group is designated group 1. A sub-group consisting of thirty-
six teaching and office staff members attending courses in the fall of
1979 is designated group 2.

Courses

Business communication

The business communication course emphasizes the transactional
writing that occurs in business and organizational settings. As a com-
munication course, it stresses clear and concise style and positive
tone. All papers are graded on the criteria of content, organization,
style, tone and mechanics. The standard of mechanics is set by the
department and allows no more than one serious error to every one hundred
words. Errors are identified by analogy from twelve violations of spell-
ing and rules of grammar and punctuation published in a student manual.
The standard is strictly adhered to in order to prepare students for a
business world that is conservative in its compliance with the conven-
tions of standard English.

Teachers do not use class time to teach mechanics. After an initial
review at the beginning of a quarter, they do discuss errors which have
occurred frequently in assignments. Papers are clearly marked for me-
chanical errors and students required to make corrections on their
papers. Final grades are based primarily on writing ability, although
there is variance among teachers in the emphasis put on quizzes. Some

teachers allow students to rewrite papers for a better grade, especially at the beginning of the quarter. Most teachers lower grades for late papers. A course manual sets minimum requirements, but no rigid standards of grading are imposed on teachers. This study attempted no controls on grading.

Teachers   The staff teaching this course specializes in this area of composition. Nine teachers taught thirteen sections during spring quarter. All teachers included in the study had taught the course at Iowa State at least twice, and a number had taught the course for more than five years. In all multi-section courses at Iowa State, an attempt is made to maintain uniformity in course design throughout the sections. Although teachers are free to draw up their own syllabi and create their own assignments, they are asked to include at least ten to twelve graded pieces of writing each quarter. Assignments are given letter grades.

Business communication short course

The six-hour business communication short course provides instruction on the principles of business writing and reviews rules of grammar. Although students do no actual writing, content, organization, tone, and style of business letters are discussed briefly in the course. Most of the class time is devoted to lecture and discussions of grammar. The course is given on consecutive days, three hours each day. Four teachers, all members of the business communication staff, conducted the short courses. Three teachers had participated in previous courses, and one teacher, who joined the group at the third session, used the teaching materials of the other members so that all sessions were presented uni-

formly.

Instruments

Business communication test

The business communication test was developed from a pool of questions submitted by the twelve members of the staff teaching the course in 1976. Teachers were asked to submit six to eight sentences that represented errors they commonly found in students' papers. A committee designed the first test by converting the sentences into 115 questions organized under four subtests: 1) spelling, 2) punctuation, 3) grammar, and 4) style and tone. The first three subtests require students to identify the statement as correct or incorrect. The last section is multiple choice. All of the questions are rhetorically representative of business letters and grammatically representative of errors typically made by this group of students.

After the initial administration of the test, items showing low correlation with total test score were re-examined, and the test was reduced to one hundred questions. The spelling subtest consists of a mock business letter with thirty-five words students must identify as right or wrong. Some words, such as "personnel," "maintenance," "employees," "accommodate," "truly," and "sincerely," are words used commonly in business. Others, such as "privileged," "separate," "definite," "argument," and "benefit," are words frequently misspelled by college students.

The punctuation subtest includes twenty-five sentences illustrating the use of comma splice, colon with a list, apostrophe, comma in re-

strictive and nonrestrictive elements, hyphen, and semicolon. The following sentences illustrate the items on this subtest:

a. Your salaried clerk need not spend his time watching for shoplifters, instead he can sell your appliances.

b. The explanation, that the credit manager offered Mrs. Smith, did not satisfy her.

The twenty-five items of the grammar section include dangling and misplaced modifiers, subject-verb agreement, pronoun case and agreement, and parallel construction. The following items are representative of this subtest:

a. Warm and cozy, the January winds will never touch you in our Alaskan Windcutter jacket.

b. Customers like Salem Feeds because they are nutritious, filling, and a promoter of growth.

The last section contains fifteen questions which ask students to identify sentences showing conciseness of style and audience-directed tone. The following items are representative of this subtest:

a. Which of these questions best shows that the writer has edited her question, eliminating unnecessary words?

1. Please send me an operating instruction booklet for the Charles Dryer, Model AX.

2. Would you be so kind as to tell me how to operate the Charles Dryer for ultimate efficiency?

3. Now that I've received the Charles Dryer, would you be able to send me the instruction booklet which was omitted from the package?

4. Would it be possible for you to explain to me how to operate the Charles Hair Dryer I bought from you last week?

b. Which sentence best emphasizes the reader?

1. Conversion of the Charles Hair Dryer into a handy portable model is possible.

2. Take your Charles Dryer with you when you travel—it easily converts into a handy portable model.

3. Charles Professional Hair Dryers have been used successfully by millions of customers.

Scoring   The test is machine scored. A score of 85 had been set as the cutoff score. Those students receiving at least 85 on the test were allowed to write an essay as the second phase of the testout procedure. A grade of B was required on the essay for the student to be given testout credit for the course. The cutoff score remained largely untested because too few students had requested permission to attempt the testout.

Reliability   Kuder-Richardson formula 20 was used to gain an estimate of the test's reliability. In the past the test had yielded coefficients of .76 to .82.

Shortened test

The shortened test used in the short courses was developed for this study and consisted of twenty-eight items drawn from the full business communication test. Students were given fifteen minutes to complete the test. Questions were chosen with three criteria in mind. The first

concern was that they represent the four sections of the original test

(spelling, punctuation, grammar, and style and tone). Second, because

the purpose of the test was to learn if the more difficult items were

learnable and thus would show gains, the questions were chosen with a

level of difficulty in the .50 range as indicated on the computer print-

out; that is, all questions had been answered correctly by students

about fifty percent of the time. Finally, questions correlating above

.20 with the total test were chosen. These questions were considered to

contribute adequately to the homogeneity of the test, as suggested by

Nunnally (1970).

For the first group the same test was administered as pretest and

posttest. Students were given fifteen minutes to complete the test.

The test was machine scored, and the standard deviation of the test was

1.91. The small dispersion of scores indicates that scores clustered

at the same level. Since adequate reliability is based in part on a

spread in scores, the reliability of the posttest was low (.04). The

inference was made that students remembered the questions from the first

test. The posttest was therefore rewritten for the second group. Ques-

tions were rephrased with care taken to present the identical grammatical

construction. The item reading, "This companies policy is explained

in the enclosed brochure," was changed in the alternate form of the

test to, "The ladies coats are featured in the latest ad." Spelling

words were not altered, but the sequence of the subtests was changed.

The second group received the original form of the pretest one month

later.

Questionnaires

Business communication questionnaire

Questionnaires were administered to students in business communi-
cation classes at the beginning and end of the quarter. The first
questionnaire had two aims: 1) to gain information on students' academic
aptitude and family educational background in order to investigate how
these effects contributed to test score, 2) to gain data on students'
confidence and motivation in grammar and writing.

The questionnaire asked students for permission to obtain ACT scores
through the registrar's office. ACT composite scores are defined as
measures of developed academic aptitude and correlate .72 with general
intelligence tests (ACT 1973). A limitation of this score is that it
was obtained in a student's junior year of high school, yet aptitude
scores are known to show relative stability (Mehrens and Lehmann 1978).
In order to measure family educational background, one question asked
students to mark their parents' highest level of education in one of
three categories: eighth grade, high school, and college or beyond.
Students also filled in cumulative grade point and high school size, but
these items were not used in the study. Most of the questions on the
questionnaire were designed to determine students' confidence in writing
and motivation for improving their writing. The questions used five-
point Likert-type scales. One question asked students to rate their
abilities in eight areas: spelling, punctuation, grammar, organization,
content, sentence structure, conciseness, and appropriate choice of
words. Another question asked students to rate these eight areas on

importance to their personal goals. Other questions asked if students

enjoyed writing and if they felt that they would be required to write

in their future professions. Permission to conduct the study and

approval of the questionnaire had been obtained from the University

Human Subjects Review Committee. Students were given an explanation

of the study and asked to sign a permission slip. (Questionnaire and

cover letter are given in Appendix A.)

Short course questionnaires

Participants received a questionnaire at the end of the short course

designed to measure their confidence and motivation and to obtain evalu-

ations of the course. Two questions asked if they composed letters them-

selves and if they edited for spelling and punctuation. Another question

used a Likert-type scale to obtain ratings of confidence on seven areas

of the course: spelling, punctuation, grammar, sentence structure, con-

ciseness, letter organization, and audience awareness. Other questions

were aimed at learning the level of education and time of high school

graduation of the adult students. The question seeking an evaluation of

the course asked students to judge the seven areas of course content as

unhelpful, somewhat helpful, and very helpful. A space for comments was

provided at the bottom of the questionnaire. An explanation of the study

and permission slip were attached to the questionnaire (see Appendix A).

Data Collection

Business communication group

Nine teachers administered the objective test to their business com-

munication classes in the first week of the spring quarter of 1979. All teachers allowed fifty minutes for the test and explained to students that they would not receive scores until the end of the quarter. Some teachers used the next class period for a review of grammar, which is customary at the beginning of a quarter, but they did not use questions from the test for discussion. The test was machine graded. Because the next class period included instruction in grammar, no makeup tests were given.

Two weeks later, teachers administered a questionnaire and asked students to sign a release in compliance with policies of the research on human subjects committee. At the time of collection teachers placed the questionnaires in envelopes and sealed them. Administering the questionnaire was delayed for two weeks because questions were included on audience, style and tone, and conciseness, and students beginning the course might have little understanding of such terms. The administration of the test and questionnaire on separate days caused a loss of subjects from absences. No makeup questionnaires were given to ensure confidentiality.

Forms allowing permission to use ACT scores were sent to the registrar's office where this information was released to the study. One hundred sixty ACT composite scores and 138 ACT English scores were available.

During the last week of the course, students retook the objective test. Teachers submitted final grades with plus and minus signs to allow for a finer discrimination in the correlations. Letter grades were then converted to numbers in the following way: 1(F), 2(D-),

3(D), 4(D+), 5(C-), 6(C), 7(C+), 8(B-), 9(B), 10(B+), 11(A-), 12(A).

Essay grading

Graded final essays were collected from four classes taught by three teachers (one had taught two classes) who volunteered their participation in grading. Since students are randomly assigned to sections, no bias occurred in student selection. All names were removed and papers coded.

Topics   While the topics for the essays had been different in each teacher's class, they all required persuasive writing. One topic was a claim letter written to a travel agency. One assignment was an employee evaluation on a new instructor, and one assignment was a job application letter. Complete details for each letter had been given to students on dittoed sheets, and the time allotted for writing was fifty minutes. (Complete assignments are given in Appendix B.)

Essay grading   To keep the grading session short and avoid rater fatigue, teachers drew forty papers at random from the combined four classes. The teachers formulated the grading criteria before the session began. Five categories were chosen: 1) content, 2) organization, 3) tone, 4) writing style, and 5) mechanics. Papers would not be graded on the Iowa State English department standard of mechanics which would have failed papers with multiple errors, but instead mechanics was given equal emphasis with the other four categories. The grading method was holistic, and no markings were made on the papers.

Teachers graded papers from the sections other than their own and marked letter grades with plus and minus signs on a class list given them. At the end of the grading session teachers regraded five papers

from their own classes to check the reliability of their original mark-
ing. Because the regrading closely matched the original, the regrading
was excluded in the reliability computation. All papers received three
grades: the original grade awarded at the time of the exam by the class-
room teachers and the grades assigned by two other teachers at the grad-
ing session. Because one of the teachers was not available for grading
in the replication, the averages of grades of two raters were used in
the final study. Essay grades were converted to the numerical scores
given above for course grade. Thirty-five of the essays were matched
with posttest scores for the correlations. Five students had been absent
at the time of the posttest.

Replication study

The replication of the study was carried out the following fall to
substantiate the reliability and predictive power of the objective test.
Because the number of essays graded in the original study was small
(forty), the addition of two sets of correlations of essay grades and
objective test scores could supply further data on the correspondence
of objective test scores and essay grades. The correspondence between
objective test and course grade could also be rechecked. The instruments
and procedure in the replication duplicated those in the major study.

The replication used three sections of the business communication
course taught by the two teachers previously rating essays. Fifty-
eight students were matched with pretest-posttest scores, but one stu-
dent, as a graduating senior, was not required to write the final essay.
His first essay was therefore not used. The teachers repeated the same

techniques for rating two essays, one at the beginning, and one at the end of the course. Essay 1 was written the first week of the course during the class period immediately following the administration of the pretest. Essay 2 consisted of the final exam in the course written one week after the administration of the posttest. Both teachers used the same topics. Essay 1 required students to write a letter describing a company they would enjoy working for. Essay 2 required students to write an informative letter explaininc a change in company policy. The second essay was more difficult to write because it involved negative information but was representative of requirements at the end of the course. Students were given fifty minutes for each assignment. (Complete assignments are given in Appendix B.) After the papers were collected, all names were removed, and papers were coded. Fifty-seven essays were graded at each session and matched with posttest scores.

Short course

At the beginning of the first session of the short course, students received the twenty-eight-item objective test. They were not informed that they would retake the test, and answers were not discussed. On the second and final day of the course, the posttest and questionnaire were administered. Group 1 was given the same test as a pretest and posttest. Group 2 received the alternate form of the test as a posttest. To determine the permanence of any gains made on the test, the study included a second posttest administered one month later. These tests were delivered to the students at their campus offices and picked up the next day, so the testing was not done under controlled conditions. The stu-

dents were asked to allot no more than fifteen minutes to the test. Two students were unavailable for retesting because they were ill. A total of thirty-four tests were collected.

Analysis

Prediction

The first objective of this study was to examine the predictive validity of the business communication objective test by correlating test scores and the criterion measures of essay grade and course grade. The Pearson product-moment correlation formula was used for both correlations. Because course grades and aptitude test scores may have a common association of students' general academic ability, which is influenced by family educational background, their correlation may be unrealistically high. To measure the association with academic ability, this study correlated test scores with ACT composite scores. In addition, chi square was used to measure the association of the test with family educational background.

The second objective was to examine the cutoff score chosen for the objective test and calculate two types of placement errors: 1) the exclusion of students from phase two of the testout procedure (writing an essay) who write well but cannot score well on an objective test of mechanics and style, 2) the inclusion of students in phase two of the testout procedure who score well on the objective test but cannot write well. Scatterplots underlying the correlations of posttest and essay grade and posttest and course grade provided evidence for both placement errors, which were given in percentages. Regression analysis was

used to estimate average essay grade and course grade for a particular cutoff score.

Content

The third objective of this study was to examine the relevance of the content of the test to the objectives of the business communication course. First, as a test of the predictive ability of each subtest, scores on the subtests were correlated with the criterion measures. Intercorrelations of subtests were calculated to measure if the subtests overlapped. Additional evidence was supplied by observing gain scores from pretest to posttest. To ensure the grammar principles were learnable in a brief time, this study used a t-test to compare gains made by business communication students, who are taught grammar indirectly in a college course, with gains made by short course students, who are taught grammar directly. Finally, scores on the objective test were correlated with scores on the ACT English test to learn if the two tests measured similar skills.

CHAPTER IV

REPORT OF FINDINGS

Three objectives for this study were set in Chapter I.

1. To examine the correlation of the business communication objective
   test with the criterion measures of essay grade and course grade.

2. To examine the data for the purpose of establishing a cutoff score
   that will minimize two types of errors of placement: 1) the exclu-
   sion of students from phase two of the testout procedure (writing
   an essay) who can write well but cannot score well on an ob-
   jective test of mechanics and style, 2) the inclusion of students in
   phase two of the testout procedure who score well on the objective
   test but cannot write well.

3. To examine if the content of the objective test fairly represents
   the course and students.

This chapter will present the results of the statistical analyses
used to accomplish these objectives. The first section gives the de-
scriptive statistics of the objective test including means, standard devi-
ations, and reliability coefficients, and the reliabilities of the cri-
terion measures of essay grade and course grade. The second section
presents the correlation coefficients of the objective test and the cri-
terion measures. Scatterplots underlying these correlations provide
evidence for setting appropriate cutoff scores and determining the per-
centage of "placement error." The third section analyzes the test's
content. Included are the correlation coefficients of the subtests and
the criterion measures and an analysis of the gains made from pretest

to posttest.

Descriptive Statistics of the Objective Test and Criterion Measures

Determining objective test scores at an intermediate level of difficulty

The purpose of the objective test is to identify high performing students as those with potential of testing out of the business communication course. Therefore, a requisite of the test is that it spread scores out over a wide range and show a fine discrimination of scores at the upper levels. If the test were so easy that a large number of students scored 100, it would not select better students. On the other hand, if the test were so hard that even good students would be forced to guess at the answers, the test would also fail to identify good students.

A hypothetical test can serve as an example of determining a mean appropriate for the business communication test. If a one hundred-item test required students to fill in an original answer for each item, guessing might yield an expected score of zero. If, on the other hand, the one hundred items were true-false, guessing would yield a score of approximately 50. Therefore, a score of 50 would be expected on a one hundred-item true-false test which was so hard that essentially everyone was forced to guess at the answers. Similarly, the expected score on a multiple choice test on which every item allowed three options and on which students guessed at every item would be 33.3. The business communication test contains eighty-five true-false items and fifteen multiple choice items (eleven with three options and four with four options). The expected score achieved by guessing on this test is 47.17.

If the test were to yield an average score this low, it would not discriminate among students' performances. Neither would it discriminate among performances if the average score were near 100. However, an average score halfway between the two extremes could be expected to show a meaningful spread of scores. For this test, a score halfway between guessing and complete mastery of the test is the average of 47.17 and 100, or 73.58. A mean of approximately 73.58 would indicate an intermediate level of difficulty. In addition, test results should show relatively few perfect scores to indicate that good students were able to demonstrate their abilities without being hampered by a test that was too easy. A mean of 73.58 accompanied by the occurrence of few perfect scores would indicate that the test shows a spread of scores which should lead to a good discrimination.

<u>Means</u> <u>of</u> <u>full</u> <u>test</u> <u>and</u> <u>subtests</u>   Tables 2 and 3 give a description of the number of items and format of the full pretest and posttest and subtests in the major study and replication. In addition, they give actual means and calculated intermediate levels of difficulty. The intermediate score between guessing and complete mastery of the full test is 73.58. The pretest mean is 73.42 in the major study and 73.97 in the replication. Posttest means, as expected, are somewhat higher. No students scored 100 on either test. Therefore, the test fulfills the scoring criteria set for its purpose of screening.

Observation of the subtests' actual means and calculated intermediate scores offers information on their respective levels of difficulty. The means of the spelling subtest (28.17-28.81) are slightly above the intermediate level (26.25); however, on the posttest, which

shows the higher means, only five students received a perfect score in the major study and one student received a perfect score in the replication. Therefore, this subtest shows an appropriate level of difficulty.

The means of the punctuation subtest (18.42-19.59) are very close to the calculated intermediate score (18.75). Six students received a perfect score in the major study, and one received a perfect score in the replication. Therefore, this subtest shows an appropriate level of difficulty.

The means of the grammar subtest (15.57-16.38) are below the calculated intermediate score (18.75). The data show that no students received a perfect score in either the major study or replication. This subtest appears to be somewhat difficult.

On the other hand, the style and tone subtest shows higher means (11.01-12.60) than the calculated intermediate score (9.83). In addition, twenty-six students received a perfect score on the posttest in the major study, and five received a perfect score on the posttest in the replication. This subtest is easy for students and may not discriminate scores in the upper levels.

Table 2. Actual means and intermediate scores for objective test and
          subtests in major study

Test	No. of items	Format	Mean (pretest)	Mean (posttest)	Intermediate scores
n = 258 pretest					
n = 208 posttest					
full test	100	true-false/ multiple choice	73.42	76.27	73.58
spelling	35	true-false	28.26	28.81	26.25
punctuation	25	true-false	18.42	19.32	18.75
grammar	25	true-false	15.73	15.70	18.75
style and tone	15	multiple choice	11.01	12.45	9.83

Table 3. Actual means and intermediate scores for objective test and
          subtests in replication

Test	No. of items	Format	Mean (pretest)	Mean (posttest)	Intermediate scores
n = 58					
full test	100	true-false/ multiple choice	73.97	77.02	73.58
spelling	35	true-false	28.17	28.45	26.25
punctuation	25	true-false	18.72	19.59	18.75
grammar	25	true-false	15.57	16.38	18.75
style and tone	15	multiple choice	11.52	12.60	9.83

Frequency distributions

Statistical calculations performed on data in testing have been de-
signed under the presumption that the numbers have a frequency distribu-
tion that is normal (Mehrens and Lehmann 1978). Normality assumes a
single mode with a symmetrical distribution of scores with 68% falling
within one standard deviation of the mode and 95% falling within 1.645
standard deviations of the mode. If the normality assumption is not
approximately met by the data, these percentages need no longer apply,
and statistical descriptions of data no longer can be encapsulated into
mean and standard deviation. The normality assumption pertains to the
large hypothetical population from which the data are a sample. The
data are not expected to be precisely normally distributed, though it
is expected that the larger data sets will be more nearly normally dis-
tributed. The frequency bar charts (Figures 6 and 7, Appendix C) for
the pretest and posttest in the major study exhibit the characteristics
of a normal distribution. The deviations shown by the bar charts in the
replication (Figures 8 and 9, Appendix C) are typical of data for smaller
sample sizes.

Measures of dispersion for the objective test

Standard deviation and variance   Many of the assessments of an
instrument such as the business communication objective test involve
measures of dispersion. The most common of these is the overall standard
deviation of test and subtest scores or the square of the standard devi-
ation, which is called variance. The standard deviation reflects in part
the meaningful spread of scores made possible by an intermediate level

58

of difficulty of the test, as discussed in the above section (Mehrens and Lehmann 1978). Standard deviation also reflects a spread other than the one due to true differences in ability. This spread is called measurement error and is caused by variable effects not related to true ability, such as factors pertaining to the test and the circumstances under which it was taken which lead to a student's answer not reflecting what he actually knows. Tables 4 and 5 give the standard deviations of the pretest and posttest and subtests in the major study and replication. "One standard deviation" designates the distance 68% of all scores will fall from the mean. Since the standard deviation of the pretest in the major study is 8.70, 68% of all scores will fall within 8.70 points of the mean. It is clear from the tables that of the subtests, spelling shows the greatest spread of scores (3.46-4.15). The style and tone subtest in the posttest of the replication shows a distinctly low standard deviation (1.64). A discernible reason for the low standard deviation is that because the test is easy, scores have clustered at a high level.

Reliability of objective test

An important measure related to dispersion is a test's reliability. Kuder-Richardson formula 20 is commonly used to estimate reliability and was used in this study. The formula is as follows:

$$r_{11} = \frac{k}{k-1} \left(\frac{1-\Sigma pq}{\sigma_x^2} \right)$$

r_{11} = reliability

k = number of items

p = proportion of persons answering correctly

q = 1 minus p

σ_x^2 = variance of the test

In this formula, variance denotes the total test variance. Summation pq is the summation over all test items of the observed proportion of students answering correctly on that item multipled by the observed proportion answering incorrectly. As such, it estimates the sum of the variances, reflecting spread of all types of all items of the test. The coefficient r_{11} has two roles. In its first role it gives an approximation of the correlation between the given test and a hypothetical similar test. Such a hypothetical test is defined to be a test on which items correlate with one another at the same level at which the items of the current test intercorrelate with one another. Also, it is assumed that items cross correlate with the items of the given test with the same correlation coefficient yielded between items. A second role of r_{11} is that it conservatively estimates the ratio of the meaningful portion of test variance (that not related to error) to the total test variance. From the formula it is clear that a large overall test variance will tend to induce a high measure of reliability. However, if the variance is accompanied by a large summation pq of item variances, the effect of the large test variance is neutralized.

Reliabilities of full tests and subtests are given in Tables 4 and 5. The reliabilities of the full test range from .73-.79. Coefficients of about .80 are usually given as acceptable (Nunnally 1970).

Subtests with fewer items can be expected to show lower coefficients. Of the subtests, spelling shows the highest reliability (.63-.75). Grammar shows the lowest (.18-.49). As noted above, lower standard deviations tend to be associated with lower reliabilities. For example, the low reliability of the subtest style and tone on the posttest of the replication (.33), shows a markedly low standard deviation for that subtest. Lower reliabilities of prediction tests are important primarily for their effect on the correlation of test and criterion measures (Nunnally 1970). The error increases the random component of scores and blurs the true relationship. The result is a lowering of the correlation coefficient.

Standard error of measurement

A third measure related to standard deviation is the measure of variability associated with measurement error, called standard error of measurement. Although this measure provides only an upper estimate of the standard error of measurement, it is typically used to pinpoint the error itself. The standard error of measurement designates the discrepancy between an individual's true score (the score obtained if no error occurred from variations in the test and test conditions) and the obtained score on the test (Nunnally 1970).

A true score can be expected to lie within two standard errors of measurement on either side of the obtained score. Table 4 shows the standard error of measurement of the full pretest to be 3.96; therefore, a true score of an individual will be within 7.92 points on either side of the obtained score. As explained above, the standard error is

of importance in this study primarily for its tendency to blur the correlation of the objective test and its criterion measures.

Table 4. Kuder-Richardson reliability coefficients, standard errors of measurement and standard deviations of objective test and subtests in major study

Test	Reliability r_{11}	Standard error of measurement $\sigma_{meas.}$	Standard deviation σ_x
n = 258			
full pretest	.79	3.96	8.70
spelling	.75	2.07	4.15
punctuation	.61	2.01	3.20
grammar	.49	2.15	3.00
style and tone	.59	1.58	2.46
n = 208			
full posttest	.79	3.81	8.40
spelling	.71	2.04	3.80
punctuation	.59	1.90	2.97
grammar	.40	2.17	2.79
style and tone	.55	1.36	2.03

Table 5. Kuder-Richardson reliability coefficients, standard errors of measurement and standard deviations of objective test and subtests in replication

Test	Reliability r_{11}	Standard error of measurement $\sigma_{meas.}$	Standard deviation σ_x
n = 58			
full pretest	.77	3.86	8.09
spelling	.63	2.10	3.46
punctuation	.24	1.95	2.23
grammar	.39	2.09	2.68
style and tone	.54	1.50	2.21
full posttest	.73	3.71	7.20
spelling	.74	2.05	4.02
punctuation	.33	1.85	2.27
grammar	.18	2.07	2.28
style and tone	.33	1.34	1.64

Reliability of essay grades

Reliability coefficients were calculated for three sets of graded essays. The original essay grading for the major study was done by three teachers. Each teacher submitted the grades she had awarded the paper at the time of the final examination. Then each paper was awarded a grade by the two other teachers. Correlations of the three graders were .57, .62, and .68. Because one of the three teachers was not available for grading in the replication, grades from raters one and three were used in the final computation. These raters, correlating .62, represented the median of the three correlations. The correlation was then converted using the Spearman-Brown prophecy formula as suggested by Diederich (1974): $\frac{2r}{1 + r}$. This formula gives a reliability coef-

ficient for the correlation. The final reliability of the two raters was computed as .77. Diederich gives .67 as a typical reliability for two raters; therefore, these coefficients are above average.

The replication furnished two further sets of test scores to be correlated with corresponding essays. Raters one and three rated both sets of essays. The correlation of the raters for essay 1 was .59. After the coefficient was converted using the Spearman-Brown prophecy formula, the reliability was computed as .74. The correlation for essay 2 was .63, which was converted with the Spearman-Brown prophecy formula to a reliability of .77.

Reliability of course grade

No attempt was made in this study to standardize final grading procedures beyond the general agreements of the teachers of the course. After course grades were collected, it was clear that grades were not homogeneous. This lower reliability may affect the correlations made with course grade. Percentages of course grades given by teachers in the major study are listed in Table 6. Percentages for the two teachers in the replication are listed in Table 7.

Table 6. Percentages of each letter grade given by nine teachers in major study

Teacher	Grade				
	A	B	C	D	F
1	.04	.42	.50	0	.04
2	0	.62	.33	0	.05
3 (2 classes)	.12	.38	.40	.05	.05
4 (2 classes)	.10	.40	.42	.06	.02
5	0	.40	.55	.05	0
6	.08	.46	.46	0	0
7	.13	.71	.17	0	0
8 (2 classes)	.02	.42	.42	.13	0
9 (2 classes)	0	.63	.38	0	0

Table 7. Percentages of each letter grade given by two teachers in replication

Teacher	Grade				
	A	B	C	D	F
1 (2 classes)	.03	.47	.42	.08	0
2	.05	.33	.62	0	0

Prediction

Reliability and measures of dispersion offer important information on the objective test's capability as a measurement instrument. This information answers the concern of how much the test can be counted on to provide similar data when it is used at another time for another group of business communication students. But it does not address the practical concern of how well the test serves the use it is intended, that is, predicting these students' writing ability. For the specific purpose of answering objectives one and two in this study, the correlations between the test scores and the criteria chosen to represent writing ability are crucial. These correlations and the scatterplots underlying them will show how precisely the test and grades on essays or grades in business communication classes correspond.

Pearson product-moment correlation assumptions

The strength of the relationship between the business communication objective test and its criterion measures, essay grade and course grade, may be measured by correlating test scores and essay or course grades using the Pearson product-moment correlation equation. Pearson product-moment correlation coefficients are meaningful when the following two assumptions obtain: 1) the relation of the two quantities being examined is near linear, 2) the variances of the values of one variable are approximately constant for different levels of the other variable (homoscedasticity) (Ott et al. 1978). Departure of the data from either of these two requirements typically has the effect of the correlation coefficient not reflecting the strength of an existing relationship; for

example, when there is a pronounced curvilinear relationship, the correlation coefficient as a measure of strength of the relationship would be unjustifiably low.

Correlation of test scores and essay grades

A scatterplot graphically displays the relationship of test scores and essay grades. Analysis of the three scatterplots (Figures 5-7) reveals a strong linear relationship between scores and grades in each correlation, documented by the fact that the average essay grade corresponding to a given test score would fall very nearly on a straight line. The essay grade variances are reasonably constant for different levels of test scores. Therefore, the correlations adhere to both assumptions required of Pearson product-moment correlations, and the resulting coefficients meaningfully represent the strength of the association of test scores and essay grades.

Significance probabilities   Significance probabilities are indications that the statistical phenomenon under investigation exists, that is, in the case of correlations, that some association between the two measures exists. As calculated, the probability estimates the likelihood that the association does not exist; therefore, the lower the probability, the more likely it is that some association exists. Probabilities of .05 or lower are usually taken to signify that the statistical phenomenon exists (Ott et al. 1978). Significance probabilities will be given with each correlation.

Correlation coefficients   Thirty-five graded essays written at the final examination period by the subjects in the major study were

matched with posttest scores. Essay grades were converted to numerical scores (F = 1; A = 12), and the average scores of the two raters are given on the ordinate axis of Figure 1. The test score is given on the abcissa axis. The Pearson product-moment correlation coefficient of test scores and essay grades was computed as .56 with a significance probability of .0004.

In the replication, two further correlations of test scores and essay grades were computed. Fifty-seven pretest scores were matched with grades on essays written the same week the test was administered (designated essay 1). The correlation of test scores and essay grades was .54 with a significance probability of .0001. Posttest scores and corresponding essay grades written during the final examination period (designated essay 2) were correlated and yielded a coefficient of .65 with a significance probability of .0001. Scatterplots for both correlations are given in Figures 2 and 3. The literature shows a range of correlation coefficients for objective tests and written essays from .42-.85 (Nisbet 1955; Diederich 1974; Johnson 1976; ETS 1979). The coefficients in this study fall within the range.
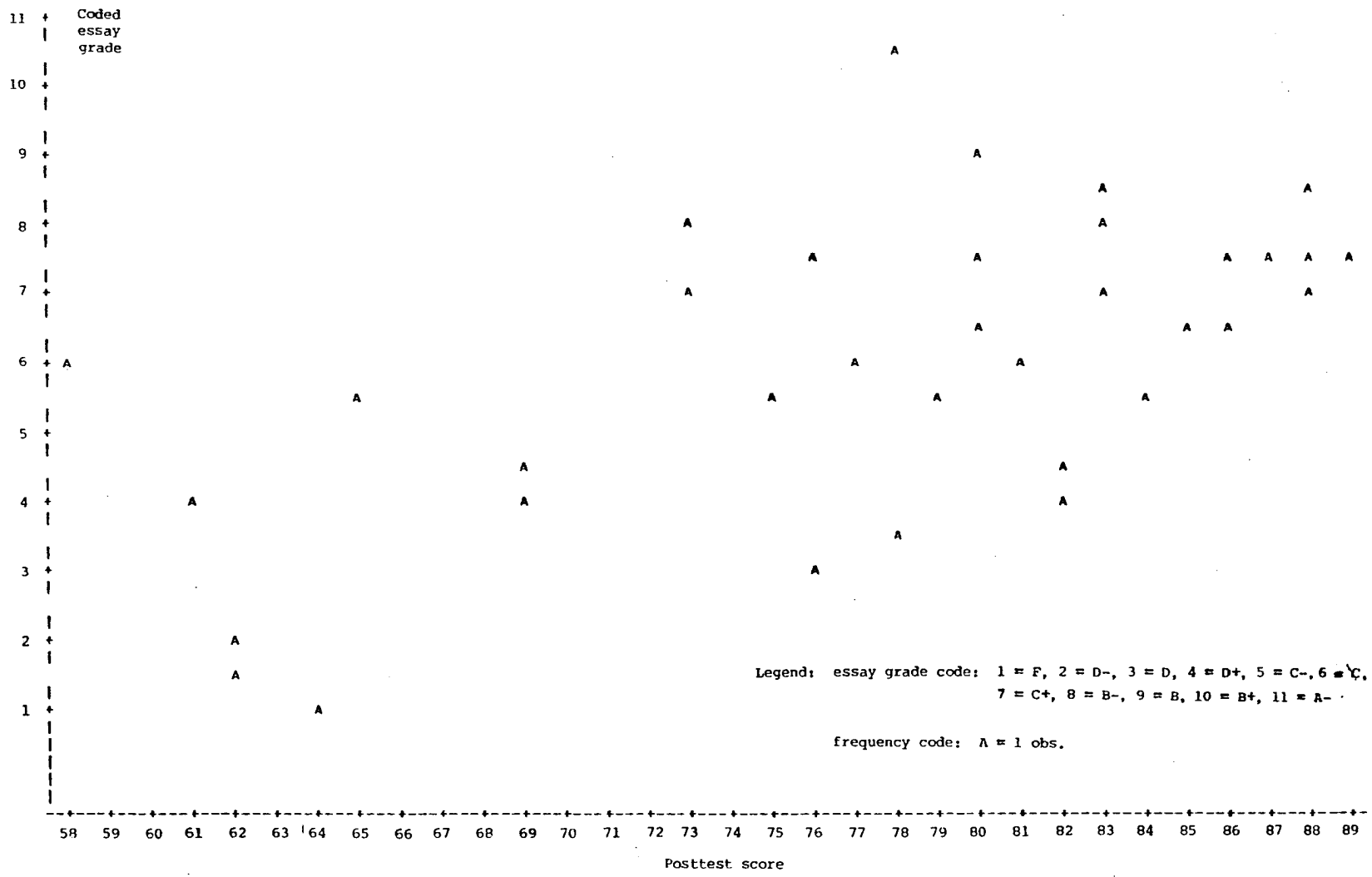
Figure 1. Scatterplot of essay grade and posttest score in major study

10 + Coded
| essay
| grade
9 +
|

8 +
|

7 +
|

6 +
|

5 +
|

4 +
|

3 +
|

2 +
|

1 +

Legend: essay grade code: 1 = F, 2 = D-, 3 = D, 4 = D+, 5 = C-, 6 = C,
7 = C+, 8 = B-, 9 = B, 10 = B+

frequency code: A = 1 obs., B = 2 obs., C = 3 obs.

50 52 54 56 58 60 62 64 66 68 70 72 74 76 78 80 82 84 86 88

Pretest score

69

Figure 2. Scatterplot of essay grade (essay 1) and pretest score in replication

Coded
essay
grade

Legend: essay grade code: 1 = F, 2 = D-, 3 = D, 4 = D+,
5 = C-, 6 = C, 7 = C+, 8 = B-, 9 = B, 10 = B+

frequency code: A = 1 obs., B = 2 obs.
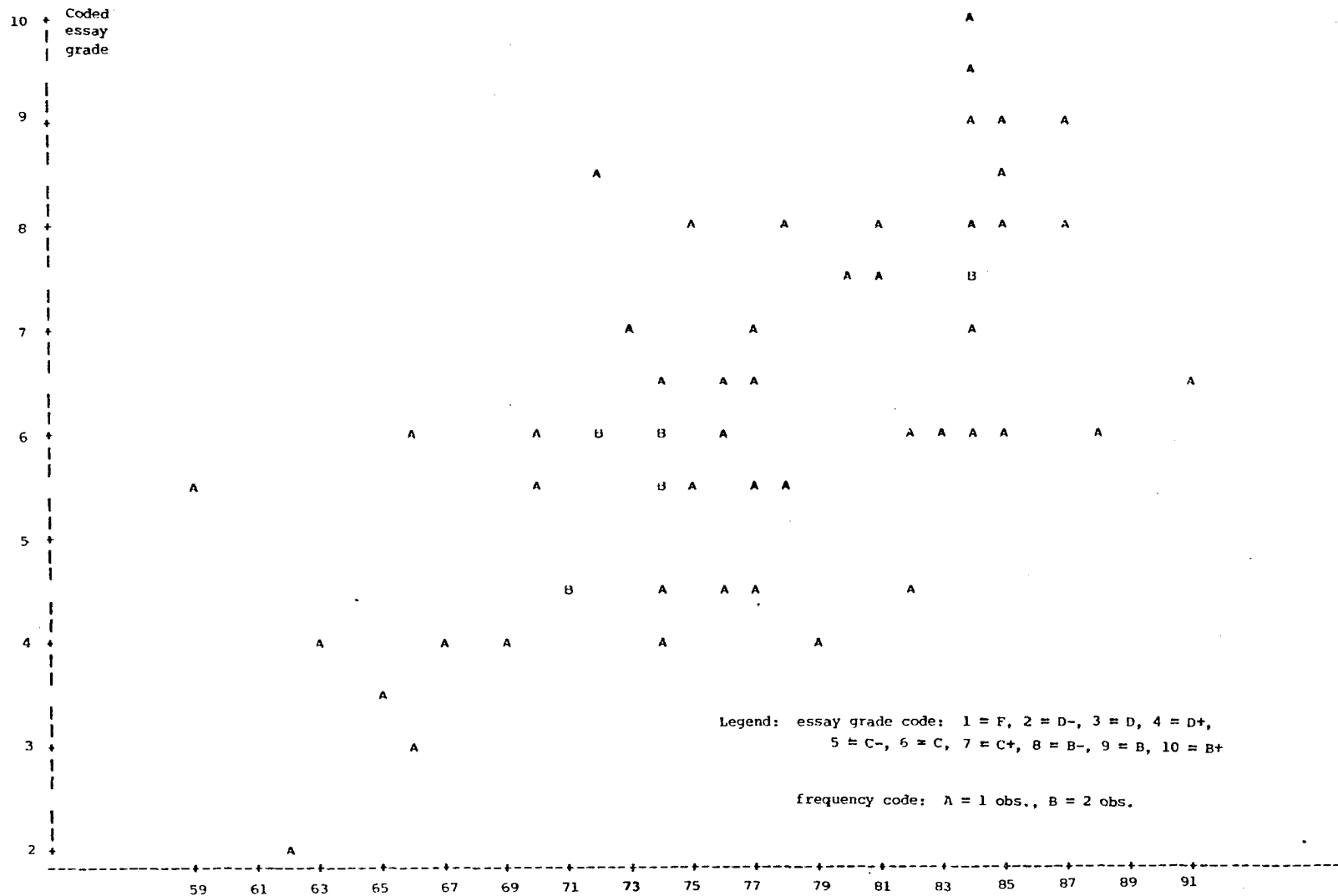
Posttest score

Figure 3. Scatterplot of essay grade (essay 2) and posttest score in replication

<u>Setting testout cutoff scores</u>   The scatterplots of the posttest

and essay in the major study (Figure 1) and posttest and essay 2 in the

replication (Figure 3) offer data for setting cutoff scores. These

scores and grades collected at the end of the course represent what stu-

dents wishing to test out of the course must achieve. In the major study,

only thirty-four of 208 students taking the posttest scored 85 or above.

In the replication, nine of fifty-eight students taking the posttest

scored 85 or above. Therefore, a score of 85 would allow only 16% of

the student population to advance to phase two of the testout procedure.

Since teachers award an essay grade of B or better to approximately 20%

of the students and a course grade of B or better to approximately 40%-

50% of the students, a score of 85 is too high. Therefore, scores of

80 and 75 will be analyzed for their ability to discriminate students

who write at a B or above level and thus would be allowed to test out

of the course.

For any score set as a cutoff below which students cannot qualify

for the second stage of the testout procedure, two types of errors can

occur: 1) students will not be identified who can write essays graded

B or better but cannot score 75 or 80 or above on the posttest, 2) stu-

dents will be identified who score 75 or 80 or above on the posttest

but will not pass the second stage of the testout procedure because

they cannot write essays graded B or better.

Error 1 results in students who are good writers being dropped from

the testout procedure. Table 8 shows the placement of students receiving

essay grades of B or above for test scores of 75 and above and 80 and

above in the major study and replication. The placement error, even for

the higher score of 80, is at 14%-25%, relatively low.

Table 8. Placement of students graded B or better on written essay for scores of 75 and 80 or above on the objective test in major study and replication

Study	Test score	No. students receiving B or above	Placed correctly	Placed incorrectly	Placement error
major	75 or above	4	4	0	0%
replication	74 or above	7	6	1	14%
major	80 or above	4	3	1	25%
replication	80 or above	7	6	1	14%

Table 9 shows the number of students who cannot write at a level of B or above and constitute error 2 and those who can write at this level and are correctly placed. Figures are given for scores of both 75 and 80 or above. The table shows that for students attaining these scores, 73%-84% would not write well enough to test out. The larger error of this type will mean that teachers must grade a number of essays that will not pass the standard of a B grade.

Table 9. Essay performance with placement error of students receiving objective test scores of 75 and 80 or above in major study and replication

Study	Test score	No. receiving score	No. receiving essay grade B or above	No. receiving essay grade below B	Placement error
major	75 or above	25	4	21	84%
replication	75 or above	34	6	28	82%
major	80 or above	18	3	15	83%
replication	80 or above	22	6	16	73%

As expected, the lower cutoff score of 75 will include more students who write well but have difficulty scoring well on an objective test and decrease error 1. At the same time, the lower score will include more students who cannot write an essay at the B level, and therefore increase error 2. Because error 2 tends to be high in this study, the higher cutoff score of 80 would be the better choice because it reduces this error.

Tables 10 and 11 give full information on the distribution of essay grades for scores of 75 and 80.

Table 10. Essay grade distribution in percent for students scoring 75 or above on the posttest in major study and replication

	A	A-	B+	B	B-	C+	C	C-	D+	D	D-	F
major study (25 scores)	0	4	0	12	28	20	20	4	8	4	0	0
replication (34 scores)	0	0	6	12	29	15	26	9	3	0	0	0

Table 11. Essay grade distribution in percent for students scoring 80 or above on the posttest in major study and replication

	A	A-	B+	B	B-	C+	C	C-	D+	D	D-	F
major study (18 scores)	0	0	0	16.6	33.3	28	11	5.5	5.5	0	0	0
replication (22 scores)	0	0	9	18	36	9	23	5	0	0	0	0

Prediction of essay grade from posttest score   A regression of essay grade on posttest score makes possible a prediction of essay grade based on a given posttest score. The prediction will show the average

grade for a certain posttest score. The formula is as follows:

$$g = a + (b)(po)$$

where

	major study	replication (essay 2)
g = essay grade		
a = intercept =	-5.40	-5.70
b = slope =	.147	.155
po = posttest score		

A posttest score of 75 predicts an essay numerical score of 5.7 in the major study and 5.9 in the replication, or an essay grade of C. A posttest score of 80 predicts an essay numerical score of 6.4 or essay grade of C in the major study and numerical score of 6.7 or essay grade of C+ in the replication. Since the relatively high test score of 80 predicts an essay grade of C+ at best, the regression explains why error 2 is high. A test score of approximately 90 would be required to predict a grade of B. Calibrations of score and essay grade must be set by departments keeping in mind fairness to students and efficiency of grading. In addition, the standards set for students to achieve in order to test out of the course will affect the balance. Lowering the essay standard to B- would decrease error 2 dramatically (about 30 percentage points) without raising error 1 unusually high (about 10 percentage points).

Correlation of test scores and course grades

The criterion measure of course grade may also be used to check the predictive validity of the test. Because of the variability of the standards used for awarding final grades, course grade may be a less

useful criterion measure than graded essay. The outcome of using a less reliable criterion would normally be a lower correlation coefficient for test scores and final grades. However, because high test scores and high course grades may have the common association of general academic ability, the correlation may be artificially elevated.

Bias of objective test and general academic ability   A correlation of scores on the business communication test and scores on a test of general academic ability would indicate if the two types of tests are closely related. To test for this association, this study correlated pretest scores with students' ACT composite scores. ACT composite score is defined as a measure of developed academic ability (ACT 1973) and represents the average of four subtests given in English, math, social science, and science. The pretest score was used in this correlation because the students' common experience of taking the business communication course would influence the posttest scores. In the major study, a Pearson product-moment correlation coefficient of .34 with a significance probability of .0001 was computed for pretest scores and ACT composite scores using 165 subjects. The low correlation suggests that the test is not biased by a strong association with general academic ability. The replication confirms this result. Pretest and ACT composite for students correlated .40 with a significance probability of .0146. The literature gives correlations of .58-.60 for tests of English aptitude and tests of general intelligence (Starbuck 1933; Stalnaker 1933; ACT 1973).

Bias of objective test and family educational background   Because all aptitudes are developed in part by environment, students not exposed

to environments that foster academic values may receive lower scores. Family educational background may influence how well students perform on aptitude tests, such as the business communication objective test, and on course grade. To check the association of students' family educational background and test score, this study issued a questionnaire asking students to mark their family education background as grade school, high school, and college or above. The category of grade school was not used because it yielded too few responses. The chi square test of independence was used to identify whether an association exists between test scores and family educational background.

To be meaningful, chi square tests must fulfill two requirements: 1) each cell must have an adequate number so that calculated significance levels will be valid, 2) continuous data cannot be grouped so coarsely as to cloud possible relationships. To fulfill both requirements, the following codings were made:

Test scores	Code
less than 65	1
65-69	2
70-73	3
74-77	4
78-80	5
above 80	6

Chi square yielded a significance probability of .2159 in the major study. No significant association exists between test scores and family educational background.

In the replication test scores were coded in the following way:

Test Scores	Code
less than 66	1
66-71	2
72-74	3
75-77	4
78-82	5
above 82	6

Chi square yielded a significance probability of .9331 and confirmed that no significant association exists between test scores and family educational background.

Correlation coefficients of test scores and course grade   Low correlations between the objective test scores and ACT scores and low associations between test scores and family educational background indicate that the test is not markedly biased. However, the correlation between test scores and course grade may be depressed by the error of measurement of course grades. An adequate correlation will indicate that course grade is a suitable criterion. Examining the scatterplots of posttests and course grades for the major study and replication (Figures 4 and 5) shows that the assumptions for the Pearson product-moment correlation are both met. Both studies show the relationships to be near linear with average grade and varying test scores falling nearly on a line. A reasonably constant variance for different levels of scores and grades obtains. Therefore, the coefficients meaningfully represent the strength of the relationships.

In the major study, the Pearson product-moment correlation coefficient for posttest and course grade was computed as .52 with a

12 +   Coded   A   A A   A
  | course
  | grade
  |
11 +   A   A   A   A   A   A A
  |
  |
10 +   A   A B   C A B A   A A B   A   A   A
  |
  |
 9 +   A   A A A   A   A A A B B D H E C F A A A B C A A C B
  |
  |
 8 +   A B   C B   A C A A B C B A C A   C A A B A
  |
  |
 7 +   A   B   A A B   A
  |
  |
 6 +   A   A A A D E B D A G C A C G B A A B A B B B B B B B   A
  |
  |
 5 +   A   A   A B   A A A   A   A A   A A   A   A
  |
  |
 4 +
  |
  | Legend: course grade code: 1 = F, 2 = D-, 3 = D,
  | 4 = D+, 5 = C-, 6 = C, 7 = C+, 8 = B-,
 3 +   A   A   A   A A   9 = B, 10 = B+, 11 = A-, 12 = A
  | frequency code: A = 1 obs., B = 2 obs.,
  | C = 3 obs., D = 4 obs., E = 5 obs.,
 2 +   A   A   F = 6 obs., G = 7 obs., H = 8 obs.
  |
  |
  |
  |
  --------------------------------+--------------------
         55 57 59 61 63 65 67 69 71 73 75 77 79 81 83 85 87 89 91 93 95 97

                                     Posttest score

Figure 4. Scatterplot of course grade and posttest score in major study

78

Figure 5. Scatterplot of course grade and posttest score in replication

significance probability at the .0001 level. In the replication, the

coefficient was computed as .57 with a significance probability of

.0001. The literature shows coefficients of from .40 to .50 for such

correlations; therefore, these coefficients are somewhat high. Course

grade seems to represent a suitable criterion measure for the test

and confirms the findings of the correlations of essay grades and test

scores. Correlations for criterion measures of essay grade and course

grade are given in Table 12.

Table 12. Pearson product-moment correlation coefficients and signifi-
cance probabilities for objective test scores and criterion
measures of essay grade and course grade in major study and
replication

Criterion measure	Number	Correlation with objective test scores	Significance probabilities
Post essay--major study	35	.56	.0004
Essay 1--replication	57	.54	.0001
Essay 2--replication	57	.65	.0001
Course grade--major study	208	.52	.0001
Course grade--replication	58	.57	.0001

Setting testout cutoff score using course grade   The scatterplots

of test scores and course grade can serve in a similar way to those of

test and essay grade in setting a cutoff score. It is important to note

that essay grade and course grade are not based on the same criteria.

Course grade includes not only writing ability but also class participa-

tion and objective test scores. Beyond these criteria is the subjective

estimate of the teacher in determining a final grade. The result of the discrepancy in the criteria used for essay and grade is that test scores may not correspond to the same grade levels for each.

Examination of the scatterplots of posttests in the major study and replication and corresponding course grades (Figures 4 and 5) shows the percentage of error of the two types: 1) students who are not identified by the objective test scores of 75 and 80 or above who can achieve grades of B or above in the course, 2) students who are identified by scores of 75 and 80 or above on the objective test but cannot achieve B in the course. Percentages of errors of placement of both types may be calculated for test scores of 75 and 80 or above.

Table 13 shows the number of B or better students identified by scores of 75 and 80 in the major study and replication. The error represents the students who would write B or better but cannot score 75 or above on the test (error 1). The percentage of placement error is 16%-36%.

Table 14 shows the number of students who write B or better essays and those who cannot write at a B or better level. Figures are given for cutoff scores of 75 and 80 or above. This placement error (error 2) is somewhat higher, 30%-51%.

Table 13. Placement of students graded B or better on course grade for
scores of 75 and 80 or above on the objective test in major
study and replication

Study	Test score	No. students receiving B or above	Placed correctly	Placed incorrectly	Placement error
major	75 or above	82	69	13	16%
replication	75 or above	22	17	5	23%
major	80 or above	82	59	23	28%
replication	80 or above	22	14	88	36%

Table 14. Course grade performance with placement error of students
receiving objective test scores of 75 and 80 or above in
major study and replication

Study	Test score	No. receiving score	No. receiving course grade B or above	No. receiving course grade below B	Placement error
major	75 or above	121	69	52	43%
replication	75 or above	35	17	18	51%
major	80 or above	84	59	25	30%
replication	80 or above	23	14	9	39%

The higher score of 80 identifies more students receiving a final
grade of B or above and represents a better cutoff score. Grade distri-
butions for students scoring 75 and above and 80 and above in the major
study and replication are given in Tables 15 and 16.

Table 15. Course grade distribution in percent for students scoring 75 or above on the posttest in major study and replication

Study	No. receiving test score of 75 or above	A	A-	B+	B	B-	C+	C	C-	D+	D	D-
major	121	3	3	12	39	17	4	15	5	0	2	0
replication	35	6	0	20	23	11	14	26	0	0	0	0

Table 16. Course grade distribution in percent for students scoring 80 or above on the posttest in major study and replication

Study	No. receiving test score of 80 or above	A	A-	B+	B	B-	C+	C	C-	D+	D	D-
major	84	5	5	17	44	14	0	13	2	0	0	0
replication	23	9	0	26	26	4	17	17	0	0	0	0

Prediction of course grade from objective test score   Just as

average essay grade could be predicted from test score by using the

regression formula, so average course grade may be predicted in the same

way. The formula is as follows:

$$g = a + (b)(po)$$

		major study	replication
g	= course grade		
a	= intercept =	-1.77	-5.70
b	= slope =	.123	.17
po	= posttest score		

A score of 75 predicts a course grade numerical score of 7-7 1/2 or C+.

A score of 80 predicts a numerical score of 7.9-8 or B-. A comparison

of these predictions with those made by regressing essay grades on post-test scores indicates that the final course grade predicted by a given posttest score is approximately a half grade higher than the essay grade prediction; thus, students receiving B- on a single essay are likely to receive a final course grade of B. Thus, the grading standard for single essays is more stringent than that for course grade. The looser grading standard used for course grade reduces error 2 about 30 percentage points and at the same time raises error 1 about 10 percentage points. If administrators wish to test out more students, they might consider lowering the standard for essay grade to B-.

Analysis of Content

Subject areas of objective tests of composition

Because objective tests cannot easily measure the larger elements of composition, such as content and organization, they often do not examine the full scope of writing, but rather test editing skills, mainly in the subject area of mechanics (usage). The National Council of Teachers of English has urged the inclusion of items which measure other areas of writing (Purves et al. 1975). However, such subtests will offer problems of item format and limit the range of items which can be covered on a test. For instance, a subtest on paragraph organization requires longer items for students to read and thus limits the time devoted to other areas. A further deterrent to including such subtests is their failure to show high correlation with criterion measures (Noyes 1963; Johnson 1976). On the other hand, usage has shown consistently good correlation with writing samples (Diederich 1974).

The business communication objective test is used to screen students for testout in the course. Students passing this screening will then write an essay for the final judgment of their ability to meet the standards of the course. The essay is the achievement test. Yet screening tests, especially when they are used for academic courses, may predict more accurately when they approximate an achievement test (Mehrens and Lehmann 1978). When they closely conform to the course objectives, they are fairer to students. The question arises: what subtests will correlate well with the criterion measures and approximate achievement in a writing course?

Assignments in business communication are graded on content, organization, tone, style, and mechanics. Because content and organization are difficult to test, tone, style, and mechanics are more suitable choices. The business communication test contains eighty-five items of mechanics and fifteen items of style and tone. This emphasis on mechanics is unjustified if the categories of the grading criteria are given equal emphasis in the course. But they are not. Single errors in mechanics lower grades more dramatically than single errors in style and tone. The exact role mechanics plays in writing is undetermined; therefore, it may be that the emphasis on mechanics is justifiable. Additional support for its dominant role is that all items were submitted by teachers familiar with the course and students. Therefore, the principles tested are closely allied to the course and population. Finally, since mechanics has traditionally been a good predictor of writing success, it might be unwise to substitute poorer predictors. If the mechanics subtest correlates more highly with he criterion measures than

does the subtest of style and tone, then it may be wiser to include a disproportionate number of mechanics items. The most important statistics will be the correlation coefficient between subtest scores and the criterion measures.

Correlation of subtest score and criterion measures of essay grade and course grade

The business communication objective test contains four subtests: spelling, punctuation, grammar, and style and tone. Correlations of the subtests and criterion measures can show the strength of the relationship of each subject area and the chosen measures of writing ability, essay grade and course grade. These correlations for both the major study and replication are given in Tables 17 and 18. A feature of the correlations is the high coefficient for the style and tone subtest with both criterion measures. The exception is the correlation of .20 with course grade (Table 17). An explanation for this low figure may be a clustering of scores on the test. This subtest is easy for students who have taken the course, and in the smaller sample of the replication, the scores may have clustered at upper levels and failed to produce a linear relationship. The significance probability of .1284 indicates that the correlation may be zero. However, evidence from the major study suggests that this subtest correlates as highly with the criterion measures as the combined mechanics subtests. The literature shows a range of correlations of .23-.53 for subtests of department-designed objective tests and graded essays (Edmiston and Gingerich 1942; Johnson 1976). The Harston study (1930) gives a range of .29-.51 for correlation of subtests and course grade. The correlations in this study fall

approximately within this range.

Table 17. Pearson product-moment correlations and significance probabilities of posttest subtests and essay grade in major study and replication

Subtest	No. of items	Correlation		Significance Probabilities	
		Major study	Repli- cation	Major study	Repli- cation
n = 35 in major study n = 57 in replication					
spelling	35	.43	.45	.0102	.0004
punctuation	25	.52	.40	.0013	.0019
grammar	25	.33	.52	.0537	.0001
combined mechanics	85	.54	.56	.0008	.0001
style and tone	15	.54	.44	.0007	.0006

Table 18. Pearson product-moment correlations and significance probabilities of posttest subtests and course grade in major study and replication

Subtest	No. of items	Correlation		Significance Probabilities	
		Major study	Repli- cation	Major study	Repli- cation
n = 208 in major study n = 57 in replication					
spelling	35	.30	.49	.0001	.0001
punctuation	25	.38	.36	.0001	.0051
grammar	25	.45	.44	.0001	.0006
combined mechanics	85	.47	.57	.0001	.0001
style and tone	15	.47	.20	.0001	.1284

Intercorrelations of subtests

Because writing is a complex process, subtests should measure a variety of the required skills. High intercorrelations of subtests and full test could indicate that subtests are overlapping. A subtest that intercorrelates highly with others might be dropped from the test. Table 19 shows the intercorrelations of the four subtests in the major study. Figures for the replication are given in Table 20. The figures show that none of the subtests intercorrelates with one another especially well. However, the subtest of style and tone shows decidedly lower coefficients. Significance probabilities in the replication indicate that no correlation may exist between this subtest and the subtests of spelling and punctuation. The data indicate that style and tone may examine different skills from mechanics subtests.

Table 19. Intercorrelations of subtests* in major study

Subtest	Spelling	Punctuation	Grammar	Style and Tone
n = 208				
spelling	xx	.44	.32	.36
punctuation	.44	xx	.47	.30
grammar	.32	.47	xx	.29
style and tone	.36	.30	.29	xx

*All significance probabilities are .0001.

Table 20. Intercorrelations of subtests with significance probabilities in replication

Subtest	Spelling	Punctuation	Grammar	Style and Tone
n = 57				
spelling	xx	.41 (.0013)	.39 (.0028)	.19 (.1583)
punctuation	.41 (.0013)	xx	.28 (.0338)	.04 (.7526)
grammar	.39 (.0028)	.28 (.0338)	xx	.34 (.0090)
style and tone	.19 (.1583)	.04 (.7526)	.34 (.0090)	xx

Pretest-posttest gain scores

Another method of checking if the content areas of a test are relevant to the course is to measure the gains made in scores from pretest to posttest (Nunnally 1970). Gains for the full test and subtests for college students in the major study and replication are given in Table 21. The total gain of the combined mechanics subtests in the major study is 1.41 and in the replication is 2.05. On the other hand, the style and tone subtest gains 1.44, or 51% of the total gain on the test. In the replication, the gain of 1.09 on the style and tone subtest represents 35% of the total gain on the test. This higher gain could be expected because style and tone are taught directly in the course, but gains in mechanics are especially low.

Table 21. Mean gains from pretest to posttest in full test and subtests with standard deviations and significance probabilities in major study and replication

Test	Major study			Replication		
	Mean gain	Std. dev.	Sig. prob.	Mean gain	Std. dev.	Sig. prob.
n = 208 in major study n = 58 in replication						
full test	2.85	5.54	.0001	3.14	5.00	.0001
spelling	.54	2.50	.0019	.40	2.68	.2648
punctuation	.89	2.67	.0001	.84	2.16	.0042
grammar	-.03	2.76	.8805	.81	2.67	.0245
combined mechanics	1.41	4.81	.0001	2.05	4.55	.0011
style and tone	1.44	2.17	.0001	1.09	2.23	.0005

Shortcourse gain scores   While it is clear that gains made in

mechanics from pretest to posttest are small, the cause may be that gram-

mar is not taught directly in the course or it is possible that grammar

principles are too difficult to learn in a single college quarter. To

test if the principles of grammar required in the course are learnable

in a short time, this study included students in short courses where

grammar is taught directly. Gains made by both college students and

short course students may be compared. This comparison is not made with

the assumption that the college students and short course students are

homogeneous groups. (See Appendix D for a description of short course

students.) The intent is to show that when these grammar principles

are directly taught, they are not so difficult that students cannot

learn them.

The shortened test was made up of twenty-eight questions taken from

the full test. The short course students are divided into two groups,

and gains from pretest to posttest are given in Table 22. (Descriptive
statistics of the shortened test used in this study are given in Appendix
E.) The groups make significant gains; however, taking the test on con-
secutive days may allow students to remember the items. Group 2 was
tested again one month later, and further gains occurred. These gains
cannot be solely attributed to the course, but they do indicate that
learning is not extinguished within a month's time.

Table 22. Mean gains from pretest to posttest made by short course
students on 28-item shortened test with standard deviations
and significance probabilities

	Number	Gain from pretest to posttest	Standard deviations	Significance probabilities
group 1	33	2.15	2.32	.0001
group 2	36	1.55	3.02	.0039
group 2 (one month later)	34	2.09	3.45	.0013

<u>Comparison of gains for college students and short course groups</u>

It is possible by using the computer to select the twenty-eight questions
used on the shortened test for the short course groups from the full test
taken by the college students. In this way the two gain scores may be
compared and the difference in gains tested for significance using the
following group comparison t-test:

$$t = \frac{\bar{X} - \bar{Y}}{\sqrt{(\frac{1}{n_1} + \frac{1}{n_2}) \frac{(n_1-1)s_X^2 + (n_2-1)s_Y^2}{n_1 + n_2 - 2}}}$$

where

$\overline{X}$ = short course means

$\overline{Y}$ = college student means

s_X^2 = variance of short course gains

s_Y^2 = variance of college student gains

Table 23 gives pretest and posttest means of the college student and short course groups and the gains made by each. Short course students gain approximately 1-1 1/2 points more than college students from pretest to posttest on these twenty-eight items. This difference in gains based on the above formula is significant at the .05 level. In this study, the college students' and short course students' pretest means are nearly equal. Since both groups begin at the same level, the comparison of average gains for the two groups is meaningful.

Because the shortened test contains only twenty-eight questions, segmenting it into subtests is risky, but item analysis shows that gains are distributed over all areas of the test. From this evidence, there is no reason to believe that the principles of mechanics on the test are inaccessible to students through the teaching process. College students do not make strong gains in grammar in the course, probably because these principles are not directly taught. Consequently, from the standpoint of what content is taught and learned in the course, the high proportion of grammar items on the objective test is not justified.

Table 23. Pretest and posttest mean scores, gain scores, standard deviations, and significance probabilities for college students and short course students on 28 items of the objective test

	n	Pretest	Posttest	Gains	Standard deviations	Significance probabilities
college students	216	19.62	20.23	.62	2.58	.0023
group 1	33	19.76	21.91	2.15	2.32	.0001
group 2	36	20.28	21.83	1.55	3.02	.0039

Correlation of objective test and ACT test of English composition

As a final check on the content of the business communication test, students' scores were correlated with the ACT test of English composition. A good correlation would indicate that the test was measuring skills similar to those measured by ACT. For this study, 160 pretest scores were correlated with ACT scores and yielded a coefficient of .57, significant at the .0001 level. Although the correlation is not especially high for two aptitude tests of composition, indications are that the tests measure similar skills.

Summary

The reliability of the business communication objective test of .73-.79, while somewhat low, is adequate for a prediction instrument (Nunnally 1970). The criterion measure of essay grade showed Spearman-Brown reliability coefficients of .74-.77, which are above the average given in the literature for essay rating (Diederich 1974). Course grade showed the marked variation that might be expected when no controls are

placed on the teachers' grading.

The more important statistic for a prediction test is the correlation coefficient for these scores and the criterion measures. This correlation for test scores and written essays ranged from .54-.65, which is within the range given in the literature (Nisbet 1955; Diederich 1974; Johnson 1976; ACT 1979). Although course grade and objective test scores may have a common association of students' general academic ability and family educational background, the objective test showed low association with both factors. Therefore, course grade proved to be a satisfactory measure of writing ability. Course grade and test scores yielded a correlation of .52-.56, which is somewhat high (Starbuck 1933; ACT 1973).

A cutoff score set at 80 would produce two types of errors: 1) 14%-25% of students who could write well but could not score well on an objective test would be excluded from attempting the second phase of the testout procedure, 2) 73%-84% of students who scored well on the objective test would be included in the second phase of the procedure but would not testout. Error 2, though somewhat high, is similar to figures given in the literature (Pearson 1978). Therefore, a cutoff score of 80 would be satisfactory.

To examine the content of the test, the study correlated the subtests with the criterion measures. Although the last subtest of style and tone was the shortest, the correlation of this subtest with both criterion measures was nearly as high as the correlation with the three grammar subtests combined. Gain scores calculated for the test showed low gains for the eighty-five items of mechanics, but gain on the

fifteen-item style and tone subtest was approximately equal to that in mechanics. Higher gains shown by students in a short course where mechanics is directly taught indicated that while mechanics principles are learnable, they are not heavily integrated into the teaching and learning process of the business communication course and may be disproportionately emphasized on the test.

CHAPTER V

DISCUSSION AND CONCLUSIONS

Summary of the Study

This study was conducted to evaluate an objective test used for screening for testout in a business communication composition course. The testout procedure consists of two phases. In the first phase, students take the objective test. Those students who score the designated cutoff score then proceed to the second phase of the test, which requires a written essay. If the objective test is to continue to be used, an evaluation must show it to have acceptable correlation with criterion measures chosen to represent writing ability. Furthermore, the cutoff score must be set so that the process is fair to students and yet time-efficient to departments. Finally, the content of the test must be representative of the course objectives.

To evaluate the objective test, this study administered the test as pretest and posttest to students in business communication classes. The test scores were correlated with two criterion measures chosen to represent writing ability, essay grade and course grade. Percentages of placement error were calculated for two possible cutoff scores. The content of the test was examined in several ways. Subtests were correlated with the criterion measures to observe the predictive ability of each subject area of the test. Gain scores for each subtest were calculated in order to examine if the content of each was a part of the teaching and learning process of the course.

A discussion of the findings and listing of the conclusions follow.

Discussion of Findings

Reliability

A test must exhibit acceptable reliability for assurance that a new sample of business communication students would perform similarly on the test. The Kuder-Richardson reliability calculated as .73-.79, while somewhat low, is adequate for a prediction test.

Prediction

A suitable relationship must exist between test scores and scores on the criterion measures chosen to represent writing ability, essay grade and course grade. The correlation coefficients computed as .52-.65 for essay grade and .52-.57 for course grade are high enough to validate this relationship. Despite some variation, course grade showed good correlations with the objective test.

Cutoff score

Students who do not attain the cutoff score will be dropped from the testout procedure; therefore, this score should be set reasonably low. However, efficiency of time and constraints of resources dictate that this goal cannot be carried to the extreme. Overall, the administration and staff must decide on the total number of students that are to be allowed to test out each year. Three aims, fairness to students, fairness to staff, and the level of achievement to be met by students who are tested out, combine in a decision for setting scores. A cutoff score of 80 was selected as fair to students, though somewhat inefficient to departments. With this score, about 25% of students who could write

at a level of B or above would be excluded from phase two of the test-out procedure. About 83% of those students who could not write an essay at a level of B or above would be included in phase two of the procedure. A higher cutoff score would select fewer students to write essays and save teachers some time, but concern for fairness to students dictates keeping the first error low. A possible adjustment could be made at the second phase of the procedure. The standards for writing could could be lowered to B-, thus allowing more students to test out. Such an adjustment would increase error 1 only slightly and decrease error 2 substantially. But such decisions must be made by the staff and administration of the department.

Because the business communication objective test shows adequate predictive ability, it is a usable instrument as it stands. On the other hand, results of this study have suggested that revision would improve the test's content.

Content

Because testout in the business communication field is new, no research has been published on objective testing for this course. This test was designed using the common model or testing in English composition that emphasizes principles of grammar (called usage on objective tests). Both ACT and SAT objective tests consist of items of grammatical usage. But unlike these standardized tests, which are designed for diverse students who will enter freshman English courses taught in diverse ways, the business communication objective test uses items submitted by teachers familiar with the course and students. Furthermore,

these items are placed in a business, rather than a literary register, and closely represent the course.

Grammar   However closely related to the course and population, these items, 85% of the test, are not proportionately representative of the course. Yet, there is reason to allow grammar more dominance on the test. Errors in grammar in business writing are serious in that they undermine not only the message but the image of the company the writer represents. Grammatical errors on an essay or poem are not judged as harshly. Standards for creative writing have loosened, and some writers purposely violate conventions for effect, while no such aberrations are allowed in business. To reflect this strict standard in the business world, teachers grade weaknesses in grammar more severely than weaknesses occurring in other areas.

Although grammatical correctness is essential in the course, a level of competence is assumed for college upperclassmen, and basics of grammar are not directly taught during class time. Instead, teachers address students, who enter the course at varying levels of competency in these areas, through comments and grades on writing assignments. This study shows that while the grammatical principles are learnable, even in a short period of time, students do not make substantial gains in scores from pretest to posttest. However, data from correlations calculated in the replication suggest that changes do take place in students' grammatical skills. The correlation between pretest and essay 1 (.52) was greater at the end of the course. Posttest and essay 2 yielded a correlation of .65. This strengthening of the association

suggests that while students do not learn new principles of grammar, they sharpen their abilities to use principles they know. They learn to effectively apply their grammatical skills and carefully proofread for errors. This observation is supported by students' comments when they evaluate the course.

Despite this emphasis on correctness in grammar in the course and the improvement students show in their facility for its use, writing is not form alone. Grammar is only one of five categories on which course work is judged. Content, organization, style, and tone make up the other categories. Good writing requires good ideas and a logical development of them. However, these skills are difficult to examine on an objective test. The last two areas, style and tone, are testable and are represented by fifteen items on the test. Because these principles are directly taught in the course, this sparse representation may constitute a flaw in the test's design.

Style and tone   The subtest on style and tone correlates strongly with the criterion measures of the test, indicating that it is closely related to the writing done in the course. Furthermore, it shows good gains from pretest to posttest, reflecting the part it plays in the teaching and learning process. Its correlations with other subtests are low, and it is likely that it taps skills different from those required for grammatical proficiency. These reasons suggest that it deserves a more dominant role in the test. In addition, certain features of the course make these areas especially suitable for testing.

Writing functions   Courses in beginning composition, where most placement procedures occur, give students wide experience in

writing. James Britton (1975) has classified writing types by "function" and lists three writing functions: expressive, transactional, and poetic. Expressive writing is an account of writers' feelings, conversations with themselves or others that might be characterized as highly personal. On the other hand, transactional writing is the practical writing done in the outer world for a known audience and specific purpose of getting things done. It is largely persuasive. Poetic writing is of interest in part for the beauty of the language alone. Its audience is wide and largely unknown to the writer. Beginning writing courses ask students to experience all writing functions through personal narrative, argument, and at times, extended metaphor or poetry. Business writing, however, deals with only one language function--transactional.

Transactional writing   While writing style differs for each function, within the transactional category it is largely characterized by conciseness with an absence of figurative language or expanded vocabulary. It is clear and without embellishment. Furthermore, the writing of business demands a heightened awareness of audience in order to establish a rapport between sender and receiver that gets the business done. This rapport, or attitude of the writer, is designated tone. The subtest of style and tone of the business communication test examines these skills by asking students to identify the sentence which best achieves conciseness, clarity, or audience awareness for a business situation. Standardized tests are unable to examine styles that reflect special language functions, and they do not attempt to test tone. But departments designing a test for a

business communication course are able to capitalize on the more specific language function utilized in the course by including items testing style, audience, and tone.

Conclusions

The objectives of this study were answered in the following way:

1. The business communication test scores were correlated with essay grade and course grade and yielded coefficients of .52-.65. The relationship of the test and its criterion measures is strong enough to support the use of the test as a screening instrument.

2. Cutoff scores set at 80 will allow for two types of placement error:

   i) Approximately 25% of the students who could write essays graded B or better will be excluded from the second phase of the test.

   ii) Approximately 83% of the students who cannot write essays graded B or better will be included in the second phase of the test. The incidence of a high error 2 is less efficient for departments but fairer to students.

3. The content of the test was found disproportionately weighted to items of grammar. While grammatical knowledge is essential in the course and some items from each of the grammar subtests should remain, increasing the items on style and tone would effectively bring the test in line with the course objectives.

Suggestions for Further Study

The results of this study suggest that the subject areas included in the business communication objective test are representative of the course, but they are not proportionate to what is taught nor do they

capitalize on the special nature of the course which lends itself to objective testing. It would be useful to revise the content of this test. The choice of subject areas could be based on evidence from this study of the subtests' reliability, predictive ability, and relevance to the course. A strong proficiency in grammar is required in the course; therefore, some questions from each subtest should remain in the instrument. Since spelling is a highly reliable test and an adequate predictor, it could remain essentially unchanged except shortened from its present thirty-five items. Similarly, punctuation predicts adequately, and as an important skill, could remain in the test. On the other hand, grammar, which examines sentence structure and verb and pronoun forms, is of primary importance; however, it shows lower reliability and questionable gain in this study and thus needs major revision. It is possible that the format of the grammar subtest should be changed to multiple choice. While standard spelling and punctuation are more clearly right or wrong, and can be tested by true-false items, grammatical structures of parallelism, verb and pronoun use, and coordination and subordination lend themselves to degrees of effectiveness. They may be better tested by choosing the best of four sentences. Similarly, style and tone should remain multiple choice. The present items appear to be too easy. More difficult items should be added and this section enlarged to represent about half of the test.

Such a design, if tested with careful item analysis, should offer adequate reliability and predictive power and represent the course as closely as the framework of objective testing allows. In this way it

could present a useful model for business communication departments

needing a testout instrument.

APPENDIX A

Cover Letters and Questionnaires for
College Students and Short Course Students

To: English 302 students

From: Carol David

Subject: Research to be done on the pretest

The English 302 staff would appreciate your help in conducting a
study on the topic, "What Part Does a Knowledge of Spelling, Grammar,
and Punctuation Play in a Student's Mastery of Business and Professional
Writing?" We hope that the results will contribute to making 302 a
better course and us better teachers. In the project the date will be
used to try to answer the following questions:

1. Does a score on the pretest predict the 302 grade?
2. What components of the test are the best predictors?
3. Do students improve in their knowledge of mechanics
   by the end of the course?
4. What are the valid questions on the test?

Your participation will require that you take the pretest at the beginning
and end of the course. After the test is machine scored, you may receive
your score, but teachers will not go over the test in class. In addition,
to taking the test you will be asked to fill out a questionnaire. Your
classroom teacher will at no time see this information. In working with
the data I will code the names, and all identifiers will be removed once
the information is compiled. At no time will any individual comments or
scores be singled out.

The results of the study will be used to improve our own teaching methods,
and they may become a part of a scholarly article of dissertation with the
ultimate aim of improving the teaching of business communication.

We sincerely appreciate your help.

I agree to participate in this project and to release the accompanying in-
formation on the questionnaire for use in research for teaching, scholarly
articles or dissertations with the aim of improving the teaching of business
communication. I understand that the information will be kept confidential
and at no time will my name be used in any report or summary. I further
understand that I am participating voluntarily and that I am free to with-
draw my consent to participate at any time and the information on me will
not be used.

Signed_____Date_____

If you have any questions about these procedures, please contact Carol David,
251 Ross Hall.

Questionnaire for English 302 Students

Code Number_____ Major_____ Cumulative grade point_____ Sex_____

Please answer the following questions by circling the best response on the right.

1. Do you feel you will have to write letters and reports
   in your profession?   yes   no

2. Do you enjoy writing?   yes   no

3. What other composition courses have you taken in high
   school and college?
   1. basic composition 2. advanced composition 3. other   1   2   3

4. What was the size of your high school?
   1. 0-199 2. 200-299 3. 300-399 4. 400-499   1   2   3   4   5
   5. over 500

5. Did you take English 302 because it was required?   yes   no

6. Did you take English 302 pass-fail?   yes   no

7. May we use your final course grade in the study?   yes   no

8. May we ask the office of student records for your ACT score? yes   no

9. What was the highest grade of your parents' education?
   1. 8 2. high school 3. college or beyond   1   2   3

10. Rate your ability in the following areas.
    1. poor 2. below average 3. adequate
    4. above average 5. excellent

    spelling   1   2   3   4   5

    punctuation (commas, semicolons, apostrophes, colons)   1   2   3   4   5

    grammar (pronouns, verb agreement, parallelism)   1   2   3   4   5

    organization   1   2   3   4   5

    content   1   2   3   4   5

    sentence structure   1   2   3   4   5

    conciseness   1   2   3   4   5

    appropriate choice of words   1   2   3   4   5

11. How important to your personal goals is achievement in
    the following areas?
    1. unimportant 2. little importance 3. rather
    important 4. quite important 5. very important

    spelling   1   2   3   4   5

    punctuation (commas, semicolons, apostrophes, colons)   1   2   3   4   5

    grammar (pronouns, verb agreement, parallelism)   1   2   3   4   5

    organization   1   2   3   4   5

    content   1   2   3   4   5

    sentence structure   1   2   3   4   5

    conciseness   1   2   3   4   5

    appropriate choice of words   1   2   3   4   5

To: Participants of Business Communications Short Course

From: Carol David, Department of English, Chairperson, Business Communications

Subject: Research to be done on short courses

Our staff would appreciate your help in conducting a study on the topic,
"What Effect Do Business Communications Short Courses Have on Performance
and Attitudes of Participants." We hope that the results will contribute
to making the short courses more effective. In the project the data will
be used to try to answer the following questions:

1. Does a score on grammar, punctuation, and spelling improve
   after the course?
2. What components of the test show the most improvement?
3. Do the participants' attitudes towards writing change?
4. What kinds of attitude changes take place?

Your participation will require that you take a short test at the beginning
and end of the course. In addition to taking the test you will be asked to
fill out a questionnaire. In working with the data I will code the names,
and all identifiers will be removed once the information is compiled. At no
time will any individual comments or scores be singled out.

The results of the study will be used to improve our own teaching methods, and
they may become a part of a scholarly article or dissertation with the ultimate
aim of improving the teaching of business communication.

We sincerely appreciate your help.

I agree to participate in this project and release the accompanying information
on the questionnaire for use in research for teaching, scholarly articles or
dissertations with the aim of improving the teaching of business communication.
I understand that the information will be kept confidential and at no time
will my name be used in any report or summary. I further understand that I
am participating voluntarily and that I am free to withdraw my consent to
participate at any time and the information on me will not be used.

Signed_____Date_____

If you have any questions about these procedures, please contact Carol David,
251 Ross Hall.

Questionnaire for Business Communications

Code Number _____

Please answer the following questions by circling the best response on the right.

1. What proportion of your work involves written communication?
   1. under 25%   2. 25%-50%   3. 50%-75%   4. over 75%   1 2 3 4
2. How much of the writing do you compose yourself?
   1. under 25%   2. 25%-50%   3. 50%-75%   4. over 75%   1 2 3 4
3. Are you expected to edit for spelling, grammar, and punctuation?   yes no
4. Are you expected to edit for conciseness and style?   yes no
5. Rate your confidence in the following areas of writing.
   1. low 2. below average 3. average 4. above average
   5. high

spelling	1	2	3	4	5
punctuation (commas, semicolons, apostrophes, colons)	1	2	3	4	5
grammar (pronouns, verb agreement, dangling modifiers)	1	2	3	4	5
effective sentence structure (sentence types, active voice)	1	2	3	4	5
letter organization (positive, negative, persuasive)	1	2	3	4	5
conciseness	1	2	3	4	5
awareness of audience	1	2	3	4	5

6. Rate the degree of help you have received from the training session.
   1. not helpful 2. somewhat helpful 3. very helpful

spelling	1	2	3
punctuation (commas, semicolons, apostrophes, colons)	1	2	3
grammar (pronouns, verb agreement, dangling modifiers)	1	2	3
effective sentence structure (sentence types, active voice)	1	2	3
letter organization (positive, negative, persuasive)	1	2	3
conciseness	1	2	3
awareness of audience	1	2	3

7. Indicate the time you would prefer alloted to each topic in a 5½ hour training session.

   1 less than allotted   2. the same as allotted   3. more than allotted
   in this session   in this session   in this session

spelling	1	2	3
punctuation	1	2	3
grammar	1	2	3
effective sentences	1	2	3
letter organization	1	2	3
conciseness	1	2	3
awareness of audience	1	2	3

8. How many years of school have you attended?
   1 high school   2. one-two years college
   3. three-four years college   4. graduate school   1 2 3 4
9. Check the year of your graduation from high school.
   1. 1940's 2. 1950's 3. 1960's 4. 1970's   1 2 3 4
10. Other comments._____

_____

_____

_____

_____

APPENDIX B

Essay Assignments

<u>Assignments used in major study</u>

1. You have just returned from a tour of Europe planned by a local travel agency. The tour was an economy package which included all expenses from point of departure. Although you enjoyed the trip, two occurrences make you seek compensation on the price. First, the trip left one day late because of ground fog at the airport and you had to stay in an expensive motel overnight. Second, the Paris hotel was full and you had to stay in an out-of-the-way area. As a result you did not get to join the bus tours and you did not get your mail from home. Write a claim letter to the company asking for an adjustment you think is fair.

2. Mrs. X is a new member of the faculty. She has recently received her Ph.D. degree and has had little experience teaching. You are in charge of new teachers and you must make reports on their progress to the chair of your department. Mrs. X knows her subject well, but she has difficulty communicating with students. Her lectures are disorganized, and she delivers them in a monotone. Furthermore, she dresses in jeans and attempts to socialize with students. Although they enjoy her as a person, they are too relaxed in class and frequently disruptive. Write a memo to your department chair evaluating this teacher. The teacher will be allowed to see the memo, so be sure it is constructive.

3. You are to be a spring graduate of Iowa State University, and you

are seeking employment as a management trainee in a large firm. Your grades are average, but you have participated in a number of your dormitory's activities and been a two-year member of a professional organization at your school. You have worked during summers for a store in your home town, and throughout the course of your five-year employment, you have been able to fill in for employees on vacation; as a consequence, you have carried a good share of the responsibility at times. Write an employment application letter to the firm, which is in a nearby city, asking for an interview.

Assignments used in replication

1. (essay 1) Write a detailed description of the company you have chosen to use for this course. Assume your audience is an interested investor. Include a discussion of the product or service and its potential for sales. Explain the reasons you have chosen this business and the rewards and drawbacks of it. Give such specific details as the size, number of employees, location, profits for first year, and physical description of the office-factory. Arouse the interest of the audience.

2. (essay 2) You are employed at the national office of a large insurance company (or company of your choice). You have decided that you will add five new agents to the six agents in a large city. Profits have been good in that area, and almost doubling the staff will pay off in the long run, even if business slows for the next few months. Write a letter that will go to all of the agents explaining the

decision and inviting them to a meeting with the new staff. Give

reasons why you think the new plan will benefit them.

APPENDIX C

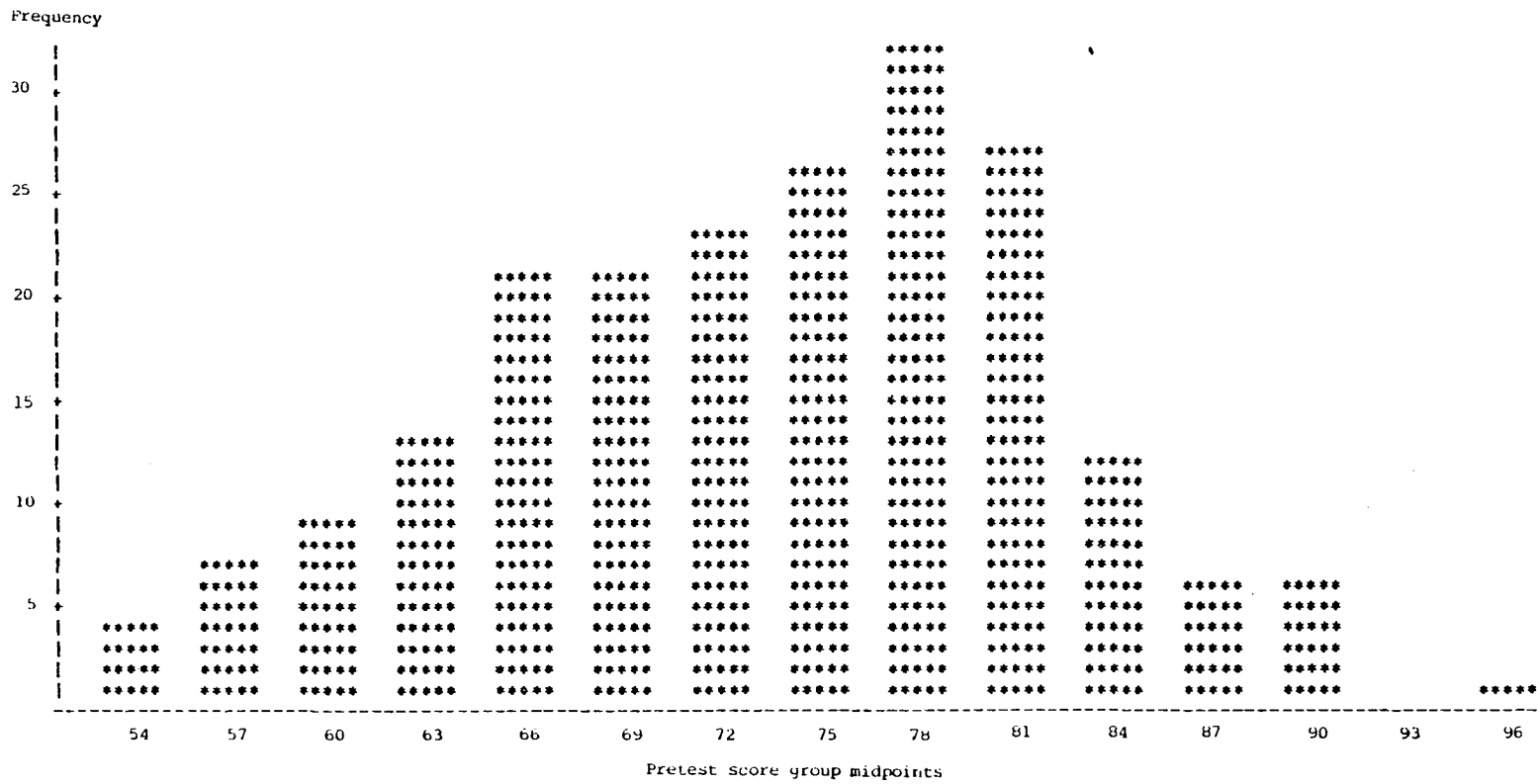Frequency Distributions of
Objective Test Scores

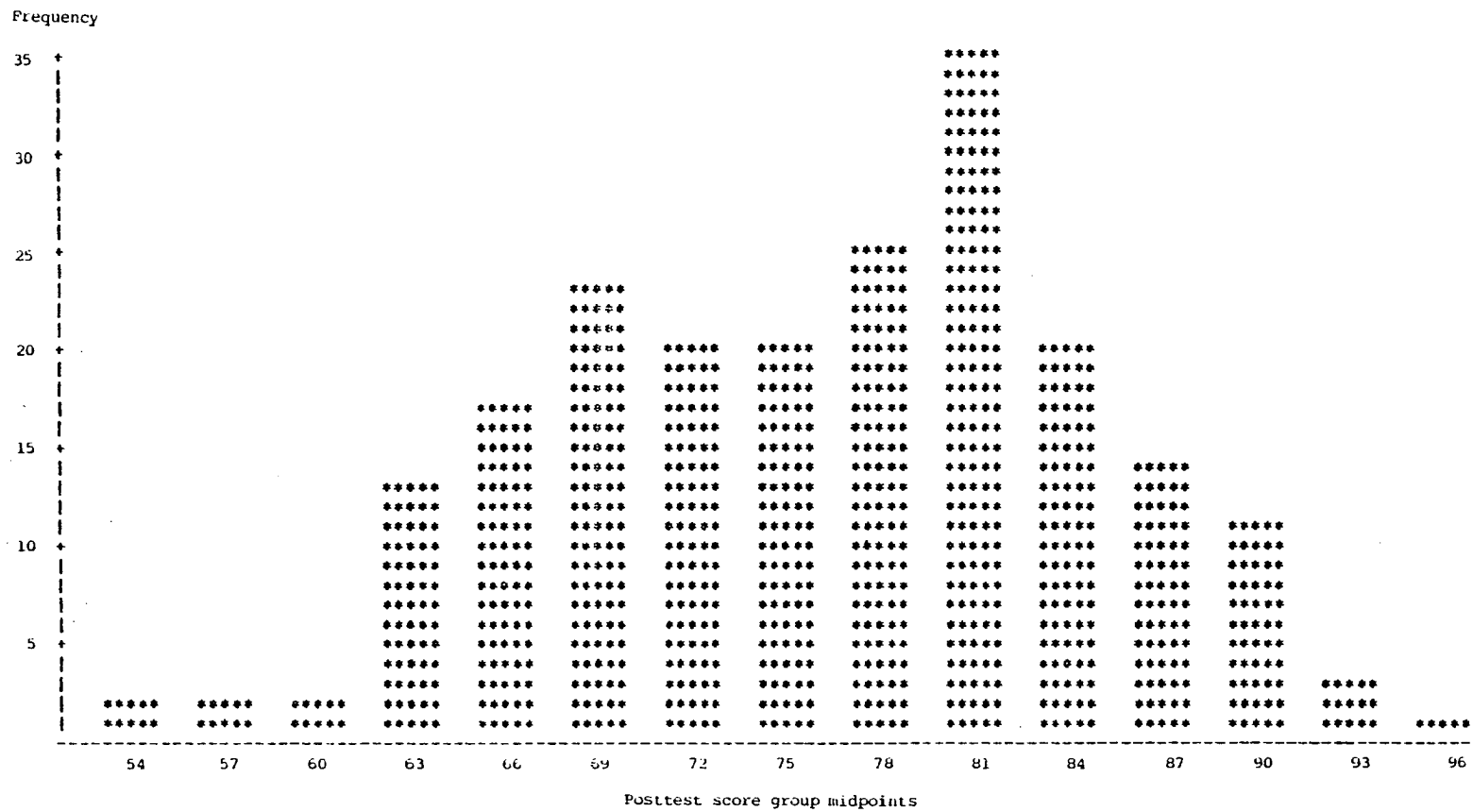Figure 6. Frequency distribution of grouped pretest score data for major study

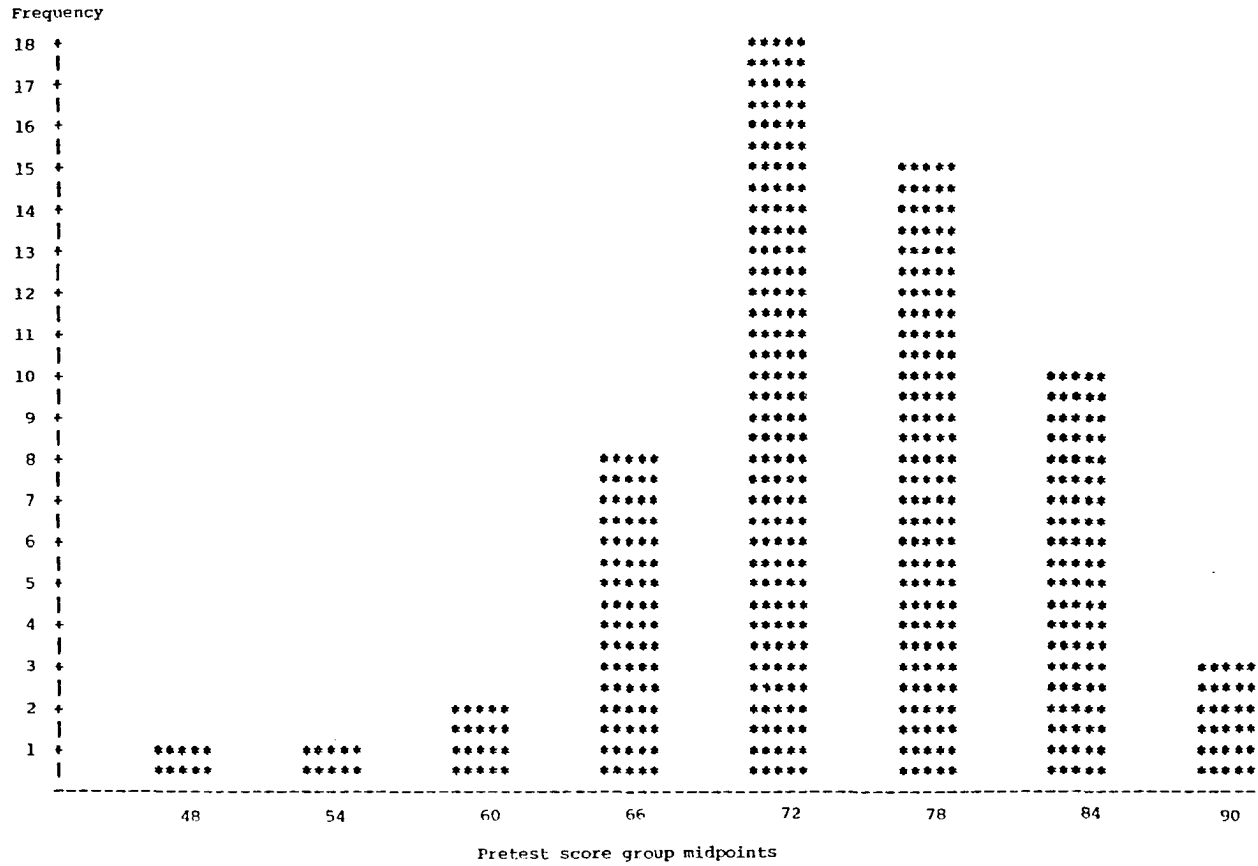Figure 7. Frequency distribution of grouped posttest score data for major study

Figure 8. Frequency distribution of grouped pretest score data for replication
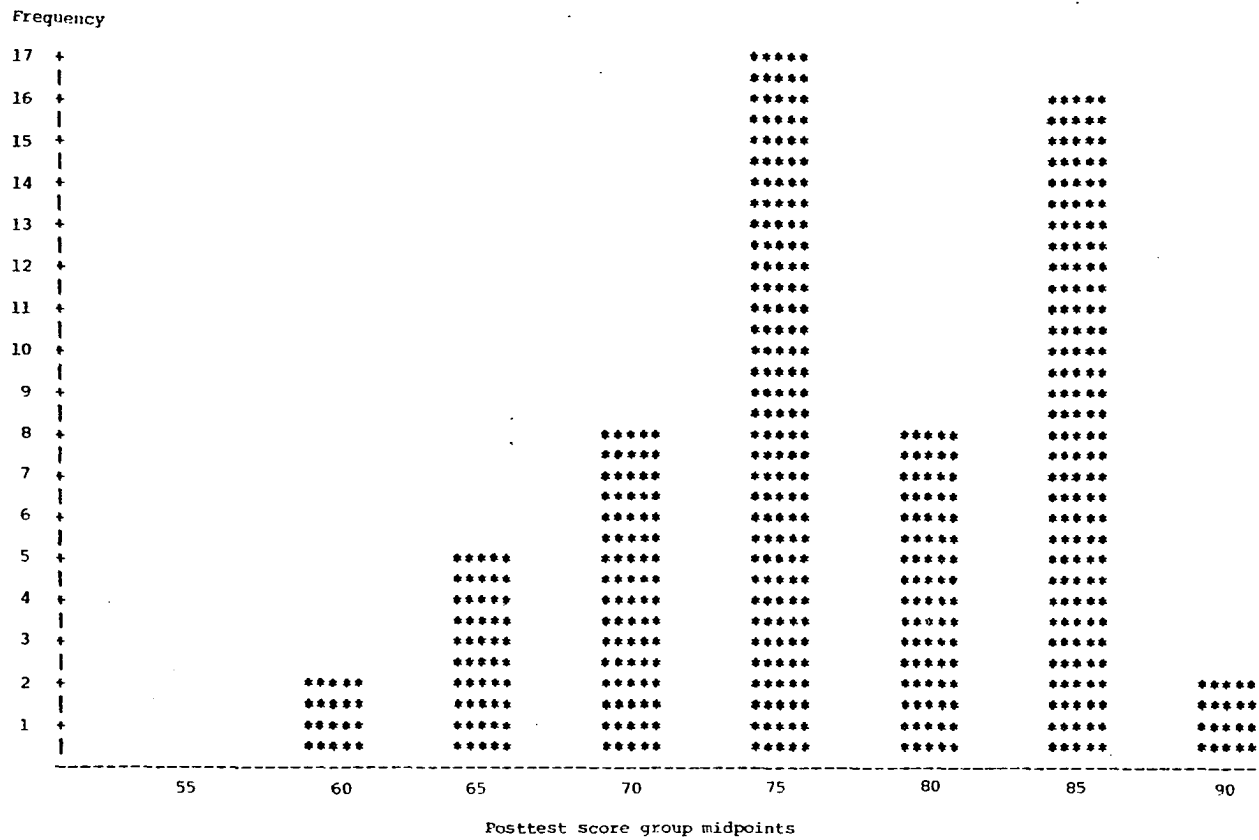
Figure 9. Frequency distribution of grouped posttest score data for replication

APPENDIX D

Description of Short Course Students

Educational background

    Data from questionnaires distributed at the short courses provide

information on the educational background and age of the students.

The mean of the educational level of the sixty-seven students was 2.6

years of college. The group was similar in educational background to

the college upperclassmen in the business communication courses.

Age

    Short course students also marked the decade they graduated from

high school. The average year of graduation was approximately 1963;

thus, students could be assumed to range in age from their late twenties

to their late thirties. The mode, however, fell in the group graduating

in the 1970's. Twenty-eight students were in this youngest group.

Confidence in subject areas

    Short course students and college students marked their confidence

in areas taught by their respective business communication courses.

Ratings for both groups on a 5-point scale are given in Table 24.

Ratings for the groups are similar.

Table 24. Means of confidence ratings given by college students and
short course students (5-point scale)

Category	Mean college students n = 208	Mean short course students n = 67
spelling	3.35	3.68
punctuation	3.13	3.48
grammar	3.18	3.27
sentence structure	3.24	3.22
organization	3.38	3.05
conciseness	3.20	3.29
Total	19.48	19.99

APPENDIX E

Descriptive Statistics of Shortened Test

Mean and reliability

The objective test administered to short course students consisted
of twenty-eight items, twenty-three true-false and five multiple choice
with three options. The expected score halfway between guessing and
perfect achievement would be 20.6. Table 25 shows that the calculated
means are very close to the expected scores; therefore, the level of
difficulty is acceptable. The reliability coefficients range from
.04-.74. The low reliability of the posttest administered to group 1
(.04) is a consequence of scores clustering at the same level. The
assumption was made that students remembered the test questions, and
the test was given in alternate form to group 2. (See Chapter III,
pp. 42-43 for a description of the test items.) A reliability of .69
was calculated for the posttest of group 2. This coefficient is
satisfactory for a short test. Means, standard deviations, reliability
coefficients, and standard errors of measurement are given in Table 25.

Table 25. Means, standard deviations, reliability coefficients and
standard errors of measurement for shortened objective test
given to two short course groups

Test	No. in group	Mean	Standard deviation	Reliability	Standard error of measurement
group 1 pretest	33	19.76	3.23	.57	2.12
group 1 posttest	33	21.91	1.91	.04	1.88
group 2 pretest	36	20.28	3.78	.70	2.08
group 2 posttest	36	21.83	3.38	.69	1.88
group 2 posttest 2	35	22.50	3.70	.74	1.88

WORKS CITED

The ACT Assessment Counselor's Handbook. 1979-80 ed. Iowa City: American College Testing Program (ACT), 1979.

Assessing Students on the Way to College: Technical Report for the ACT Assessment Program. Iowa City: American College Testing Program (ACT), 1973.

Borg, Walter R., and Gall, Meredith Damien. Educational Research: An Introduction. 3rd ed. New York: Longman, 1979.

Braddock, Richard; Lloyd-Jones, Richard; and Schoer, Lowell. Research in Written Composition. Urbana, Ill.: National Council of Teachers of English, 1963.

Britton, James; Burgess, Tone; Martin, Nancy; McLeod, Alex; and Rosen, Harold. The Development of Writing Abilities (11-18). London: Schools Council Research Studies, 1975.

Brown, Rexford. "What We Know Now and How We Could Know More about Writing Ability in America." Journal of Basic Writing 1 (spring/ summer 1978): 1-17.

Cast, B. M. D. "The Efficiency of Different Methods of Marking English Compositions." British Journal of Educational Psychology 9 (November 1939): 257-269.

Cooper, Charles R., and Odell, Lee. Evaluating Writing: Describing, Measuring, Judging. Urbana, Ill.: National Council of Teachers of English, 1977.

Diederich, Paul B. Measuring Growth in English. Urbana, Ill: National Council of Teachers of English, 1974.

Edmiston, R. W., and Gingerich, C. N. "The Relation of Factors of English Usage to Composition." Journal of Educational Research 36 (December 1942): 269-271.

The English Composition Test with Essay: A Report Prepared in Collaboration with Marjorie Kirrie. Princeton, New Jersey: Educational Testing Service (ETS), 1979.

Harston, L. D. "A Five Year Study of Objective Tests for Sectioning Courses in English Composition." Journal of Applied Psychology 14 (June 1930): 202-210.

Johnson, N. W. "The Uses of Grammatical and Rhetorical Norms, Pedagogical Strategies, and Statistical Methods in Designing and Validating a Composition Placement Instrument." Ph.D. dissertation, The Catholic University of America, 1976.

Lesikar, Raymond V. Business Communication: Theory and Application. 4th ed. Homewood, Ill.: Richard D. Irwin, 1980.

McKee, J. H. "Subjective and (or Versus) Objective." English Journal (College Edition) 23 (February 1934): 127-133.

Mehrens, William A., and Lehmann, Irvin J. Measurement and Evaluation in Education and Psychology. 2nd ed. New York: Holt, Rinehart & Winston, 1978.

Nisbet, J. D. "English Composition in Secondary School Selection." British Journal of Educational Psychology 25 (February 1955): 51-54.

Noyes, Edward S. "Essay and Objective Tests in English." College Board Review 49 (winter 1963): 7-10.

Nunnally, Jum C., Jr. Introduction to Psychological Measurement. New York: McGraw-Hill, 1970.

"Occupational Breakdown of ABCA Membership." ABCA Bulletin 43 (March 1980): 28.

Ott, Lyman; Mendenhall, William; and Larson, Richard F. Statistics: A Tool for the Social Sciences. 2nd ed. North Scituate, Mass.: Duxbury Press, 1978.

Pearson, Patricia. "A Test-Out Program." Iowa English Bulletin 28 (fall 1978): 33-34.

Purves, Alan et al. Common Sense and Testing in English. Urbana, Ill.: National Council of Teachers of English, 1975.

Stalnaker, John M. "Essay and Objective Writing Tests." English Journal (College Edition) 22 (March 1933): 217-222.

Starbuck, A. "The Freshman's First Hurdle: the Formulation, the Testing, and the Revision of a Placement Procedure." English Journal (College Edition) 22 (September 1933): 643-649.

Steele, Louise. American Business Communication Association. Urbana, Ill.: Telephone Interview. September 24, 1980.

Taking the SAT: A Guide to the Scholastic Aptitude Test and the Test of Standard Written English. Princeton, New Jersey: College Entrance Examination Board (CEEB), 1978.

Thorndike, Robert L., and Hagen, Elizabeth P. Measurement and Evaluation in Psychology and Education. 4th ed. New York: John Wiley & Sons, 1977.

Treece, Malra. Communication for Business and the Professions. Boston: Allyn & Bacon, 1978.

Warner, Joan Elizabeth. "Where Has All the Syntax Gone?" ABCA Bulletin 1 (March 1979): 17-21.

Weaver, Constance. Grammar for Teachers: Perspectives and Definitions. Urbana, Ill.: National Council of Teachers of English, 1979.

Wells, Walter. Communications in Business. 2nd ed. Belmont, Calif.: Wadsworth Publishing Company, 1977.

Wilkinson, C. W. "Grammar, Theory, and Principles in the Basic College Course in Business Communication (What? Whether? How Much?)" ABCA Bulletin 40 (September 1977): 1-4.

Wilkinson, C. W.; Clarke, Peter B.; and Wilkinson, Dorothy C. Communicating Through Letters and Reports. 7th ed. Homewood, Ill.: Richard D. Irwin, 1980.

Wilkinson, Dorothy C. "Evidence That Others Do Not Agree with Your Grading." ABCA Bulletin 42 (September 1979): 29-30.

ACKNOWLEDGMENTS