THE COMPARATIVE EFFECTIVENESS OF THE APOQT AND THE ACE IN PREDICTING ACADEMIC SUCCESS OF STUDENTS AT BOWLING GREEN STATE UNIVERSITY

Kenneth O. Johnson

Submitted in partial fulfillment of
the requirements for the degree of
Master of Arts

BOWLING GREEN STATE UNIVERSITY

BOWLING GREEN, OHIO

January, 1963

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

The incredibly rapid technological advances of the past fifty years have resulted in major revisions in the traditional American concept of military preparedness. "The stereotyped picture of the revolutionary minuteman with his rifle and his plow has for the past 185 years symbolized the military philosophy of the American people."[1] However, with the advent of inter-continental bombers, missiles, thermo-nuclear weapons, radiological and biological warfare, the concept has changed from that of the "minuteman" to that of a highly skilled group of professional technicians.

The high cost and complexity of the technical equipment in use in the Air Force today make it imperative that only those young men with a genuine aptitude for the job be selected to act as officers (supervisors) in this technically skilled group.

The Air Force Officers' Classification Manual lists more than 175 specialties for officers, which are grouped into twenty-six broad related areas.[2] There is, of course, much overlap in the aptitudes required by the different areas; however, since these job specialties range in

---

[1] Robert L. Lathrop and others, Student Attitudes Toward Reserve Officers Training Corps Programs (Minneapolis: University of Minnesota, 1962), p. 1.

[2] United States Department of the Air Force, Air Force Officer Classification Manual 36-1 (Washington: Government Printing Office, 1961), pp. 2-5.

variety from that of chaplain to jet fighter-pilot, and from accountant to space guidance engineer, there is little doubt that different aptitude patterns do exist. The continued technological advances in aerospace warfare require ever higher levels of aptitude, and make the accurate selection and classification of officers increasingly imperative. The Air Force Officer Qualifying Test (AFOQT) was designed as a selective device in an effort to meet this need.

The AFOQT was originated in 1942 as a collection of "apparatus' tests of psychomotor skills . . . used to select pilots, bombardiers, and navigators. . . ."[3] It has gradually evolved from the early mechanical test of motor skills into the pencil and paper test in use today for the ". . . selection and classification of Air Force Officers."[4] The effectiveness with which the early forms of the test predicted success in aircrew training programs led quite naturally to an expansion in use to cover officer selection and classification. The large number of men to be tested at widely scattered points throughout the nation made it impractical to continue the rather complex, mechanical, psychomotor tests.[5] Continuous validity evaluations and factor analyses led to the development of the pencil and paper test in use today.[6]

---

[3] Edwin A. Fleishman, "Psychomotor Selection Tests," Personnel Psychology, Vol. 9, No. 4 (Winter, 1956), p. 441.

[4] United States Department of the Air Force, Manual for Interpretation of Aptitude and Interest Scores of the Air Force Officer Qualifying Test (Washington: Government Printing Office, 1958), p. 1.

[5] Fleishman, "Psychomotor Selection Tests," p. 459.

[6] Ibid., p. 461.

The present form of the AFOQT consists of five aptitude scores and four interest scores. Data are drawn from three or four test booklets (depending on the form in use) to give the following composite aptitude scores: pilot aptitude composite, navigator-technical aptitude composite, officer quality aptitude composite, verbal aptitude composite, and quantitative aptitude composite.[7] This study is concerned with only the five composite aptitude scores.

I. THE PROBLEM

A tabulation of AFOQT officer quality composite scores for the 1961 freshmen AFROTC students at Bowling Green State University indicated a mean below the 35th percentile. This part of the AFOQT was designed primarily as " . . . a measure of 'intelligence' rather than of officer qualities such as leadership . . ."[8] and was, therefore, assumed by the writer to be roughly equivalent to the ACE total score.

A similar, comparative study of the ACE total percentile scores for the same students resulted in a mean of 56.22. The seemingly paradoxical disparity between the results of the two tests suggested the present study.

Purpose of the study. The purpose of this study was to compare

---

[7] Air Force, Manual for Interpretation of Aptitude and Interest Scores, p. 3.

[8] Raymond E. Christal and John D. Krumboltz, Technical Memorandum, Use of the Air Force Officer Qualifying Test in the AFROTC Selection Program, Air Force Personnel and Training Research Center (ASTIA Document Service Center, Dayton 2, Ohio, 1957), p. 35.

the relative effectiveness of the AFOQT and the ACE in predicting the academic success of selected students who have taken both tests at Bowling Green State University.

Importance of the study. The Air Force acquires the major portion of its new officers from AFROTC units located at 173 colleges and universities throughout the nation. The program, as presently constituted, has proven to be costly and ineffective in meeting the Air Force needs for new officers.[9] The current program is based on the Federal Land Grant Act of 1862 as modified by the National Defense Act of 1916. The Air Force has prepared legislative proposals for submission to the current session (1962) of Congress which would eliminate the first two years of the present program, and modify the curriculum of the second two years. If adopted, this legislation would expand the selection base from the present one of approximately 200,000 to one of about 600,000. Proposed legislation would also provide an $1100.00 scholarship fee to each enrolled student. It was anticipated that these measures would provide adequate incentive to insure a sufficient number of applicants to meet the Air Force need.[10]

Present plans call for initial testing of at least 60,000 applicants, at a great many testing centers (the exact number has not yet

---

[9] Opinion expressed by Colonel Eugene Byrne at an AFROTC staff conference, Air University, Maxwell Air Force Base, January 4, 1962.

[10] Letter from Colonel William C. Lindley, Commandant AFROTC, Air University, Maxwell Air Force Base, April 24, 1962.

been determined) throughout the nation. The present AFOQT requires trained administrators and takes approximately seven hours to complete. It is a very closely controlled device, and extensive time and effort are needed for accurate analysis and dissemination of the machine-scored results.

If it can be shown that there is a high degree of correlation between the results of the ACE and that of selected parts of the AFOQT at a large percentage of the institutions hosting AFROTC, it might be feasible to use the ACE results, or results of comparable tests, for initial selection purposes under the proposed program. If this plan should prove to be feasible, substantial savings in manpower and money could accrue to the nation.

II. DEFINITIONS OF TERMS USED

AFROTC. This term refers to the Air Force Reserve Officers' Training Corps, as organized and administered under the authority of Chapter 905, Title 10, United States Code. "Air Force ROTC is the major officer procurement program. It is conducted jointly in working partnership by the cooperating educational institutions and the department of the Air Force."[11]

ACE. The ACE is the American Council on Education Psychological Examination for college students. It is a pencil and paper test designed

---

[11] United States Department of the Air Force, Air Force Regulations 45-48, Reserve Forces (Washington: Government Printing Office, 1958), p. 1.

to measure aptitude for college academic work. It has six sub-tests which are scored to provide a total score, and two part scores, quantitative and linguistic.

AFOQT. This term refers to the Air Force Officer Qualifying Test. It is a comprehensive pencil and paper test designed to provide aptitude and interest scores for use in selecting, classifying and assigning Air Force Officers.[12]

Grade-point average. A student's grade-point average is a weighted average computed by assigning the following values to letter grades: A, four points; B, three points; C, two points; D, one point; and 0 points for each F. "A student's point average is obtained by dividing the total number of points earned by the total number of semester hours undertaken."[13]

Academic success. For the purpose of this study a student's degree of academic success was determined by his grade-point average.

Freshmen AFROTC students. This group included only those AFROTC freshmen students who took the AFOQT and for whom grade-point averages and ACE scores were available.

---

[12] Air Force, Manual for Interpretation of Aptitude and Interest Scores, p. 1.

[13] Bowling Green State University, Bulletin (Bowling Green State University Bulletins, Vol. XLVLL, No. 2., Bowling Green, Ohio, 1961), p. 38.

III. LIMITATIONS OF THIS STUDY

1. The relatively small number of Bowling Green State University male freshmen students who have taken the AFOQT tends to reduce the validity of the correlations computed and limits the generalization of the results to a larger population.

2. Some error may have resulted from the fact that no scores were available for students with grade-point averages below 1.30. This is a result of the fact that academically deficient freshmen were dropped at the end of the first semester and did not therefore, take the AFOQT, which was offered only during the second semester.

IV. ORGANIZATION OF REMAINDER OF THESIS

Literature relative to this study is reviewed and summarized in Chapter II. The procedure of the study is discussed in Chapter III. An analysis and discussion of the data collected, the statistical manipulations, and results are presented in Chapter IV. The summary, conclusions, and results are presented in Chapter IV. The summary, conclusions, and recommendations are submitted in Chapter V.

CHAPTER II

REVIEW OF PERTINENT LITERATURE

As far as can be determined there are no published reports of studies comparing the relative effectiveness of the AFOQT and the ACE in predicting academic success. Numerous related studies were found, however, in which the two tests were correlated with academic grades. A brief review of selected representative studies of this type will be presented.

The AFOQT is a controlled test that, in its present form, has not been released for public use; therefore, the studies related to it are almost entirely those prepared either by, for, or under the direct auspices of the Air Force. Because of the general lack of direct knowledge concerning the AFOQT, it will be reviewed descriptively in more detail than the ACE.

The literature relative to the ACE will be reviewed primarily as it relates directly to the problem of predicting academic success.

I. THE AFOQT

The AFOQT in use for this study was scored to obtain five aptitude scores, each of which is a composite based on a combination of nineteen sub-tests. It is a comprehensive test requiring approximately seven hours to complete. The AFOQT was standardized on a test population at ". . . approximately the officer level with respect to aptitude, and thus considerably higher than the total population of young men eligible for

military service."[1] A 1955 draft-eligible population was tested and 69.2 per cent scored below the standardized 5th percentile, while only 1.2 per cent scored above the 95th percentile.[2]

Officer aptitude. The officer aptitude composite is a compilation of scores from the following subtests: vocabulary, interpretation of data, verbal analogies, general mathematics, background for world events, reading comprehension, arithmetic reasoning, and officer biographical inventory. This composite is primarily a ". . . measure of general learning ability and officer quality."[3] Persons with high officer quality scores may be expected to do well in any course of training with high academic content and in most technical training courses with the exception of pilot training.[4]

An Air Force Academy study of 276 cadets showed a correlation coefficient of .63 between a composite of grades in six academic courses and the AFOQT officer composite.[5]

A study of two classes (1951A and 1952A) of the Officer Candidate School indicated a correlation coefficient of .40 between the officer

---

[1] United States Department of the Air Force, Manual for Interpretation of Aptitude and Interest Scores of the Air Force Officer Qualifying Test (Washington: Government Printing Office, 1958), p. 26.

[2] Ibid., p. 28.

[3] Ibid., p. 3.

[4] Ibid.

[5] Ernest C. Tupes and Raymond E. Christal, Psychological Tests and the Selection and Classification of Air Force Officers, Air Force Personnel and Training Research Center, Development Report (Lackland Air Force Base, 1957), p. 7.

composite scores and the over-all class standing. Over-all class standing included several criteria in addition to academic grades. A study of the Officer Candidate School class of 1958C, with 94 students, showed a correlation coefficient of .65 between academic grades and officer composite scores.[6] E.C. Tupes, et. al., studying an Officer Candidate School class of 221 in 1955 found a correlation coefficient of .36 between officer aptitude scores and academic grades.[7] A thorough study of the Air Force Academy class of 1959 showed a correlation coefficient of .60 between officer aptitude scores and academic grades.[8] Data from that study have been extracted, tabulated, and are presented as Table I.

Pilot aptitude. The original Army Air Corps Pilot Aptitude Test of 1942 consisted of a composite of psychomotor subtests, such as the "Complex Coordination Test," "Rudder Control Tests," and the "Rotary Pursuit," all of which were apparatus tests.[9] Validity studies of the

---

[6] Air Force, Manual for Interpretation of Aptitude and Interest Scores, p. 30.

[7] Ernest C. Tupes, A Carp, and Walter R. Borg, Validation of a Proposed Officer Effectiveness Selection Battery, Air Force Personnel and Training Research Center (Lackland Air Force Base, Texas, 1957), p. 19.

[8] John D. Krumboltz and Raymond E. Christal, Predictive Validities for First-Year Criteria at the Air Force Academy, Air Force Personnel and Training Research Center, Development Report (Lackland Air Force Base, Texas, 1957), p. 3.

[9] Edwin A. Fleishman, "Psychomotor Selection Test," Personnel Psychology, Vol. 9, No. 4 (Winter, 1956), pp. 452-4.

TABLE I

CORRELATIONS OF AFOQT COMPOSITES WITH GRADES AT THE AIR FORCE ACADEMY[10]

(1) Composite	(2) Chemistry	(3) English	(4) History	(5) Mathematics	(6) Philosophy	(7) Academic Composite	(8) N
Pilot	.23	.05	-.22	.14	-.01	.09	239
Verbal	.34	.52	.40	.20	.26	.41	239
Navigator	.46	.16	.02	.41	.18	.37	239
Quantitative	.53	.30	.28	.62	.38	.58	239
Officer	.51	.51	.43	.49	.40	.60	239

[10] Krumboltz, Predictive Validities, p. 3.

early psychomotor tests resulted in correlation coefficients of about .40 with a criterion of either graduation or elimination from pilot training. Later versions, combining the psychomotor and printed tests, reached validity coefficients as high as .70. Printed tests used alone were reported to have validity coefficients as high as .60.[11]

The present form of the pilot aptitude composite is a "measure of some of the characteristics necessary for successful completion of pilot training."[12] This composite is composed of the following subtests: pilot biographical inventory, aviation information, mechanical information, mechanical principles, visualization of maneuvers, instrument comprehension, and flight orientation.[13] Data presented in a study by E.C. Tupes and R.E. Christal indicate that the AFOQT pilot aptitude tests are ". . . good predictors of success in pilot training."[14]

Validated against the criterion of graduation versus elimination,[15] the AFOQT pilot aptitude had a validity coefficient of .42 for 947 pilot trainees during 1951.[16] An overall evaluation of more than 19,000 pilot

---

[11] Fleishman, "Psychomotor Selection Test," p. 458.

[12] Air Force, Manual for Interpretation of Aptitude and Interest Scores, p. 8.

[13] Ibid., p. 4.

[14] Tupes and Christal, Psychological Tests, p. 4.

[15] There are many factors that may cause elimination from pilot training other than pilot aptitude, e.g. sickness, injury, personal or family hardship. Therefore, this coefficient may be somewhat lower than would otherwise be warranted.

[16] Air Force, Manual for Interpretation of Aptitude and Interest Scores, p. 30.

trainees from 1949 to 1954 showed that only 11.5 per cent of the students receiving scores above the 95th percentile were eliminated while 84.2 per cent of those falling below the 5th percentile were eliminated.[17]

Though ". . . learning to fly requires a relatively high degree of intellectual capacity, . . . general intelligence tests . . . [and] . . . educational level have been found to be relatively poor predictors of pilot success. . . ."[18] Table I indicates that the pilot aptitude scores had a correlation coefficient of .09 with academic grades.

Navigator-technical. The navigator-technical aptitude composite of the AFOQT is designed to predict success in training courses for navigation, aerial observation, communications, weather, engineering, and technical intelligence, all of which have a relatively high academic content.[19] As currently constituted this composite is composed of the following subtests: interpretation of data, general mathematics, arithmetic reasoning, general science, scale reading, aerial landmarks, mechanical information, and mechanical principles.[20]

A validity study was conducted using 770 AFROTC graduates enrolled in the navigator training course during 1955-56 which resulted in a correlation coefficient of .47.[21] E.C. Tupes and R.D. Christal found a

---

[17] Tupes and Christal, Psychological Tests, p. 2.

[18] Ibid.

[19] Air Force, Manual for Interpretation of Aptitude and Interest Scores, p. 9.

[20] Ibid., p. 4.

[21] Ibid., p. 31.

". . . strong relationship between navigator-technical aptitudes and achievement. This is true for all Air Force technical schools studied. . . ."[22] Their study of grades from two technical schools, Armament and Ground Electronics, showed that of the students scoring in the top 5 per cent on the navigator-technical composite, 94 per cent received grades above the median while only 6 per cent of those in the lowest 5 per cent received grades above the median.[23] Table I shows a correlation coefficient of .37 between the navigator-technical composite and grades.

Verbal aptitude. The verbal aptitude composite is a measure of verbal skills. It was designed to predict success in such training areas as ". . . administrative services, personnel, public information, security, education and training, psychological warfare, and historical activities."[24] It is currently composed of the following subtests: vocabulary, verbal analogies, background for world events, and reading comprehension."[25] The verbal composite scores are ". . . similar to the American Council on Education Psychological Examination Linguistic scores."[26] The Air Force Academy study shows a correlation coefficient of .41 between the

---

[22] Tupes and Christal, Psychological Tests, p. 4.

[23] Ibid., p. 6.

[24] Air Force, Manual for Interpretation of Aptitude and Interest Scores, p. 11.

[25] Ibid., p. 4.

[26] Ibid., p. 11.

verbal composite and grade averages. E.C. Tupes and R.E. Christal report a "strong relationship" between the verbal aptitude scores and grades achieved in the Air Force Intelligence course, which has a relatively high verbal and general academic content.[27]

Quantitative aptitude. The quantitative aptitude composite is a measure of general mathematical ability. This composite is composed of the following subtests: interpretation of data, general mathematics, and arithmetic reasoning.[28] The score on this part of the test is ". . . similar to the American Council on Education Psychological Examination Quantitative score."[29] The quantitative composite was designed to predict success in training courses requiring ". . . mathematical and computational ability. . . ."[30] The Air Force Academy study shows a correlation coefficient of .62 between the quantitative composite and mathematics grades and .58, when correlated with grade averages.

II. THE ACE

ACE. The ACE was designed to measure "aptitude for college study."[31] It is composed of six subtests which are scored to provide

---

[27] Tupes and Christal, Psychological Tests, pp. 4-6.

[28] Air Force, Manual for Interpretation of Aptitude and Interest Scores, p. 4.

[29] Ibid., p. 12.

[30] Ibid.

[31] American Council on Education, Psychological Examination for College Freshmen (Norms Bulletin, Los Angeles: Cooperative Test Division, Educational Testing Service, 1953), Foreword.

three composite scores: quantitative, linguistic, and total.[32] The ACE test was standardized on entering college freshmen. The 1952 edition, in use at Bowling Green State University at the time of this study, was standardized on the basis of the results of 42,332 students representing 269 colleges and universities.[33]

Quantitative. The ACE quantitative composite is composed of the following subtests: (1) arithmetical reasoning, (2) number series, and (3) figure analogies. It is designed to measure the student's ". . . ability to think in quantitative terms."[34] A student receiving a high score on this test would be expected to do well in a course involving mathematics.

Linguistic. The ACE linguistic composite is composed of the following subtests: (1) same-opposite, (2) completion, and (3) verbal analogies.[35] It was designed to measure the student's general linguistic ability. Students receiving a high score on this test would be expected to do well in reading courses such as are common in the humanities. In general the linguistic test has been found to ". . . give higher

---

[32] American Council on Education, Norms Bulletin, p. 1.

[33] Ibid., p. 17.

[34] American Council on Education, Manual of Instructions (Cooperative Test Division, Educational Testing Service. Los Angeles: 4641 Hollywood Boulevard, 1950), p. 6.

[35] American Council on Education, Norms Bulletin, p. 1.

correlations with scholarship in the liberal arts colleges than do quantitative tests."[36]

ACE total. The ACE total composite score is comprised of a total of the raw linguistic and quantitative scores. It was designed to be a measure of over-all academic aptitude and would be expected to correlate positively with grades.

Predictive validity. The prediction of academic success in college has ". . . probably received more public attention than any other single problem in education. . . ."[37] The ACE was first published in 1924, and since that time there have been a great many studies designed to show its relationship to academic success. A few of the more important of these studies will be briefly reviewed in this section.

The ACE has been described as a ". . . good all-round measure of academic promise."[38] College counselors have used the ACE extensively for guidance in advising college students. L.L. Thurstone reported that 230,000 copies of the 1937, 14th edition, were ordered by 600 different educational institutions.[39] H.M. Fowler confirmed the extensive use of

---

[36] American Council on Education, Manual of Instructions, p. 2.

[37] Benjamin S. Bloom and Frank R. Peters, The Use of Academic Prediction Scales for Counseling and Selecting College Entrants (New York: The Free Press of Glencoe, Incorporated, 1961), p. 6.

[38] Albert B. Crawford and Paul S. Burnham, Forecasting College Achievement (New Haven: Yale University Press, 1946), p. 99.

[39] L.L. Thurstone and T.G. Thurstone, "The 1937 Psychological Examination for College Freshmen," Educational Record, XIX (April, 1938), 11, 209-34.

the ACE, but indicated considerable variation in the degree to which it was successful in predicting academic success. He reported validity coefficients ranging from .25 to .66.[40]

A.B. Crawford and P.S. Burnham, Yale University, credited the ACE with validity coefficients ranging from .36 to .61, but were critical because it seemed to ". . . offer little discriminating value for those institutions which maintain more rigorous standards of selection."[41] In support of this contention they reported that the Yale freshmen ". . . average score was above the year's composite 90th percentile."[42]

Ralph F. Berdie, et al., University of Minnesota, reported validity coefficients for thirty colleges ranging from a low of .13 to a high of .81 with a median of about .50, and consistently higher correlations for women than men. Their study also indicated that high school class rank correlated more closely with college grades than the ACE.[43]

Dewey B. Stuit, et al., State University of Iowa, compiled and edited a series of studies showing the effectiveness of the ACE in

---

[40] H.M. Fowler, "American Council on Education Psychological Examination for College Freshmen," The Fifth Mental, Measurements Yearbook, Oscar Krisen Buros, editor (New Jersey: The Gryphon Press, 1959), p. 427.

[41] Albert B. Crawford and Paul S. Burnham, Forecasting College Achievement (New Haven: Yale University Press, 1946), p. 92.

[42] Ibid.

[43] Ralph F. Berdie, et al., Counseling and the Use of Tests Minneapolis: University of Minnesota, 1959), p. 87.

predicting success in several "professional" schools. Eleven studies of engineering students showed correlation coefficients ranging from .21 to .55 between grades and the ACE total score.[44] Two different law school studies indicated correlation coefficients of .56 and .39 between grades and the ACE total score.[45] Nineteen studies at teacher training schools indicated correlation coefficients ranging from .37 to .62 between the ACE total score and grade averages.[46]

The Thurstones, reporting on the 1930 edition of the ACE, disclosed that 138,407 test copies were sold to 347 different educational institutions. Their report also presented validity coefficients as reported by various institutions. These coefficients are given in the accompanying table:

Carleton College	.60
Centenary College of Louisiana	.52
Colorado State Teacher's College	.49
Cornell University	.44
Duke University	.55
Syracuse University	.56

It will be noted that the correlations range from a low of .44 at Cornell to a high of .60 at Carleton, with Colorado State Teacher's College representing the approximate median of .49. In addition to those tabulated above, 19 Michigan colleges reported correlations ranging from .32 to .62; and the University of Michigan reported a series of correlations

---

[44] Dewey B. Stuit, et al., Predicting Success in Professional Schools (Menasha, Wisconsin: George Banta Publishing Company, 1949), p. 29.

[45] Ibid., p. 53.

[46] Ibid., pp. 145-6.

ranging from .348 to .560.[47]

A.H. MacPhail reported in a 1942 study the relative effectiveness of the ACE quantitative and linguistic scores in predicting grades in quantitative or linguistic type courses. He classified such courses as history, English and foreign languages as "verbal" and chemistry, mathematics, biology and psychology as quantitative. The results of his correlations indicate considerable variation in the derived coefficients as shown in the accompanying table:

	Verbal courses	Quantitative courses
ACE Linguistic score	.198 to .503	.008 to .504
ACE Quantitative score	.016 to .400	.152 to .665

He reported two examples in which the linguistic score correlated more highly with mathematics grades than did the quantitative score. In MacPhail's own words ". . . strangely enough, the Q score correlated with grades in several 'verbal' courses to a much higher extent than with grades in about half of the 'quantitative courses.'"[48]

E.E. Anderson, et al., in their study of scores made on three different psychological tests by 112 women students reported correlations between the ACE total score and grades ranging from .48 to .55. The correlation coefficient between the quantitative score and grades was .39,

---

[47] L.L. Thurstone and T.G. Thrustone, "The 1930 Psychological Examination," Educational Record, XII (April, 1931), pp. 160-78.

[48] A.H. MacPhail, "Q and L Scores on the ACE Psychological Examination," School and Society (September, 1942, p. 251.

and between the linguistic score and grades was .54. The study concluded that ". . . it is apparent that the L-score of the ACE is about as efficient in predicting general scholastic success as is the total A.C.E. score. . . ."[49] In comparison, the Revised Stanford-Binet and Wechsler-Bellevue scores showed relationships with grades at a level of .55 and .45 respectively.[50] This would indicate that the three tests were approximately equal in their predictive effectiveness.

Lee J. Cronbach summarized the results of ACE and grade correlations quite succinctly with the statement that "the total score predicts grade averages, usually with validity about .45."[51]

The above cited works are representative and, though reported correlation coefficients range from .008 to .81 the median would seem to fall between .40 and .50. In the words of Bloom and Peters ". . . it is still unusual to find a correlation between college grades and other measures above the level of .60, and most correlations reported fall in the range +.45 to +.55."[52] The disappointingly low correlations often reported may be as much the result of using ". . . such an unstable criterion as grades . . ." as from any weakness of the test itself.[53]

---

[49] E.E. Anderson, et al., "Wilson College Studies in Psychology: I.A. Comparison of the Wechsler-Bellevue, Revised Stanford-Binet, and American Council on Education Tests at the College Level," Journal of Psychology, XIV (October, 1942), p. 324.

[50] Ibid., p. 321.

[51] Lee J. Cronbach, Essentials of Psychological Testing (second edition; New York: Harper and Brothers, 1960), p. 229.

[52] Bloom, The Use of Prediction Scales for Counseling, p. 8.

[53] Crawford, Forecasting College Achievement, p. 130.

The wide variations among educational institutions with respect to aptitude level of the entering student body and the lack of standardized grading practices indicate the need for studies to establish local ". . . norms pertinent to the respective educational standards of each college. . . ."[54]

Several studies have been conducted at Bowling Green State University to establish norms and to determine the validity of the ACE as a predictor of academic success.

Margaret M. Slater's study in 1952 was designed, in part, to determine the relationship between ACE scores and grade-point averages. Her figures were based on a random sample of 400 students at different semester levels with no differentiation by sex.

Her results show correlation coefficients between grade-point averages and ACE composites as follows: total score, .492; linguistic score, .481; and quantitative score, .358.[55]

Frank C. Arnold, Director of the University Counseling Center, prepared a study of the relationship of the ACE, and other selected factors, to "progress in college." The study was a follow-up on the 1223 students enrolled as freshmen in the fall of 1955; it traced their progress in relationship to certain selected factors until June of 1959, at which time approximately 30 per cent had graduated. The entering class

---

54 Crawford, Forcasting College Achievement, p. 99.

55 Margaret M. Slater, "The Meaning of A.C.E. Scores of Students at Bowling Green State University" (unpublished Master's thesis, Bowling Green State University, Bowling Green, Ohio, 1952), pp. 18-31.

was about equally divided between men and women and showed a relatively equal distribution of scores in each decile of the ACE total score distribution. The study revealed that 39.8 per cent ranked in the upper quartile and 11.6 per cent in the bottom quartile of their high school graduating class.[56] The data presented indicate a wide range of academic aptitude, which should produce relatively high correlations between ACE scores and grade averages.[57]

His study indicated a correlation coefficient of .59 between grade-point average and the ACE total score for men students and .62 for women. The over-all correlation coefficient for all groups remained fairly stable at about .60 throughout the eight semesters studied. The slight drop in this figure from the first semester high of .62 to .59 in the eighth semester may well be explained on the basis of the loss of some lower aptitude students.[58]

Arnold's study, in line with those of other schools, found consistently higher correlations for women than for men. He also found that rank in high school graduating class had a higher correlation, .76, with grades than did the ACE total score. His conclusions indicate

---

[56] Frank C. Arnold, "The Relationship of Selected Factors to College Success at Bowling Green State University" (Bowling Green State University, Bowling Green, Ohio, 1959), pp. 1-46. (Mimeographed.)

[57] Lindquist, A First Course in Statistics, p. 195-7.

[58] Arnold, "The Relationship of Selected Factors to College Success at Bowling Green State University," p. 12.

". . . reasonable validity in predicting such aspects of college success as grades and graduating from college."[59]

III. SUMMARY

A review of the related literature reveals the AFCQT to be a rather complex selection and classification test designed to measure aptitude for pilot training, navigator or technical training and what might be termed general intelligence or learning ability. Several tests of validity indicate that the AFOQT has proven of real value when used for its designed purpose.

The ACE was designed to measure a student's aptitude for college training. A great many studies indicate that it has been a very useful device; predicting academic success with a relatively high degree of validity.

No studies were found which made a direct comparison of the AFOQT and the ACE in respect to their predictive capability. However, the literature reviewed indicates that the AFOQT officer composite should correlate with grade-point average at about the same level as the ACE total score.

In design purpose the AFOQT quantitative and verbal composites are very similar to the ACE quantitative and linguistic composites respectively. The evidence in the literature reviewed is far from conclusive,

---

[59] Arnold, "The Relationship of Selected Factors to College Success at Bowling Green State University," p. 46.

but indicates a similar range of correlation coefficients for those composites of the two tests. The AFCOT navigator-technical composite is heavily weighted with quantitative sub-tests and was found to correlate with grade-point averages in a manner similar to the ACE quantitative. The evidence indicates that the AFOQT pilot composite has little or no relationship with academic grades.

CHAPTER III

METHODS OF PROCEDURE

I. INTRODUCTION

The major purpose of this study was to compare the relative effectiveness of the AFOQT and the ACE in predicting academic success of selected students at Bowling Green State University. In order to obtain the data necessary for such a study, the following procedure was used.

II. PROCEDURE

Sample selection. The sample group used for this study consisted of the male freshmen students at Bowling Green State University who were enrolled in the AFROTC during the academic year of 1961-62. Initial enrollment in September of 1961 was 211, but 60 of these students were not enrolled during the second semester when the AFOQT was administered. Eleven new students for the second semester did not have grade-point averages comparable to the other students. Consequently the sample group used was made up of 160 AFROTC students for whom ACE, AFOQT and grade-point averages were available. The records were incomplete in some respect for 20 of these students, but were adequate for use in one or more of the comparisons made. Sample size, then, ranged from a low of 140 to a high of 153 in comparisons made. For the group used to obtain information concerning male freshmen at Bowling Green, all male students enrolled as freshmen during the first semester of the academic year of 1961-62 and

having ACE scores and grade-point averages were used.

Data collected. Student records in the Air Force ROTC office, the Counseling Center, and the Registrar's Office were used to obtain the following data: namely, AFOOT percentile scores, ACE percentile scores and grade-point average to the end of the first semester of the 1961-62 academic year.

For male freshmen in general, only the name, ACE total percentile score and grade-point average for the first semester of the 1961-62 academic year were obtained.

Plan of analysis. The relative effectiveness of the ACE and AFOOT for prediction of grade-point averages was determined by computation of Pearson product-moment correlations between appropriate pairings and comparison of these correlation coefficients. The significance of difference between correlations was determined through the use of variations of the Fisher "z" method as described by Quinn McNemar.[1]

---

[1] Quinn McNemar, Psychological Statistics (third edition; New York: John Wiley and Sons, Incorporated, 1962), pp. 130-45.

CHAPTER IV

PRESENTATION AND ANALYSIS OF DATA

The purpose of this study, as stated in Chapter I, was to compare the relative effectiveness of the AFOQT and the ACE in predicting academic success of selected students at Bowling Green State University. In order to accomplish this purpose, correlations were computed to show the degree of relationship between grade-point averages and each composite of both the ACE and the AFOQT. The observed correlation coefficients between the AFOQT composites and grade-point averages were compared with those between grade-point averages and the ACE total score and appropriate part scores. The statistical significance of observed differences between correlation coefficients was then determined in accordance with an appropriate statistical method.

The purpose of this chapter is to present and discuss the results of these statistical manipulations.

I. RELATIONSHIP BETWEEN THE AFOQT AND THE ACE

In order to determine that a valid basis existed for comparison of the AFOQT composites with those of the ACE on the effectiveness of grade prediction, it was deemed desirable to first examine their apparent relationships as stated in their design purpose.

An analysis of the data presented in Chapter II indicated some degree of similarity in purpose between the AFOQT and the ACE. The AFOQT quantitative and verbal composites appeared to correspond fairly closely

to the quantitative and linguistic composites of the ACE in their design purpose. The AFOQT officer composite appeared to correspond to the ACE total score in that the major components of both tests are quite similar and both were designed to measure general learning ability.

The AFOQT navigator-technical composite did not appear to have a direct counter-part in the ACE; however, it does include subtests designed to measure aptitude for quantitative subjects and general science. Since it was designed to measure aptitudes necessary for success in technical schools, and is composed largely of quantitative subtests it would seem to correspond in some degree at least to both the ACE total score and the ACE quantitative score.

No counterpart for the pilot aptitude composite of the AFOQT was found in the ACE. None of the subtests designed to measure "general learning ability" were included in the pilot composite since acceptance for pilot training is contingent upon successful completion of the officer composite as well.

Score distributions. To assist in the comparative analysis of the two tests, frequency distributions of scores for each test composite were set up. A frequency distribution of the ACE total percentile score for all male students of the freshmen class (1961-62) was also prepared to help determine the representativeness of the AFROTC sample group. The distribution of scores for the AFROTC sample group is presented in Table II; and that for the all male freshmen group is shown in Table III. Percentages were rounded to the nearest whole number.

TABLE II

DECILE DISTRIBUTION OF ACE AND AFOQT SCORES OF THE AFROTC SAMPLE GROUP IN JUXTAPOSITION WITH GRADE-POINT AVERAGES

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	PER CENT IN EACH DECILE								
	ACE[a]			AFOQT[b]					CRADE-POINT[c]
Decile	Q	L	T	P	V	N	Q	O	
10	21	5	12	3	5	16	16	4	1
9	21	11	11	6	10	18	10	4	8
8	15	10	14	8	8	6	14	5	9
7	7	16	14	7	7	12	10	6	18
6	7	8	15	11	9	12	8	10	37
5	10	17	8	14	15	9	12	12	20
4	6	12	9	10	12	12	9	13	7
3	7	11	9	13	16	7	10	14	
2	4	7	7	14	6	6	7	18	
1	1	3	1	14	12	2	4	14	
Total	100	100	100	100	100	100	100	100	100
Number	149	149	153	153	153	153	153	153	143
Mean	66	52	59	40	45	60	58	38	2.32
S.D.	26	24	25	32	26	26	27	25	.54
S.E.	2	2	2	3	2	2	2	2	.04

[a]Abbreviations refer to Quantitative, Linguistic, and Total percentile scores for American Council on Education Test.

[b]Abbreviations refer to Pilot, Verbal, Navigator, Quantitative, and Officer composite percentile scores on Air Force Officers' Qualifying Test.

[c]This Column divides the score range of grade-point averages in ten equal intervals from 00.0 to 4.00.

AFROTC sample group distribution. Table II indicates considerable variation in the magnitude of observed means among the different test composites. This would be expected considering the variety of factors measured and the fact that the AFOQT and ACE were standardized on different groups.

The relationship between the means of the quantitative and verbal linguistic composites was of particular interest. The quantitative mean of each test was significantly higher than was its verbal linguistic mean. These differences would seem to be in line with the published ACE norms, which show the mean quantitative score for men to be several score-points higher than that for women; while in contrast the mean linguistic score for men is several points lower than that for women.[1] This would indicate that the ACE quantitative and linguistic scores are influenced by a sex-linked factor. If this is true it might also influence the relative degree of relationship found between grade-point averages and the two composites.

All male freshmen distribution. Table III shows the mean ACE total percentile score for the all male freshmen group to be 55 and the mean grade-point average to be 2.24, neither of which is significantly different from that of the AFROTC sample group which had an ACE mean of 59 and mean grade-point average of 2.32. It should be noted that students

---

[1] American Council on Education, Psychological Examination for College Freshmen (Norms Bulletin. Los Angeles: Cooperative Test Division, Educational Testing Service, 1953), pp. 11-14.

TABLE III

DECILE DISTRIBUTION OF ACE TOTAL PERCENTILE SCORE AND GRADE-POINT AVERAGE FOR THE ALL MALE FRESHMEN GROUP

	PER CENT IN EACH DECILE	
	ACE	
Decile	Total Score	Grade-point[a]
10	9	1
9	13	5
8	12	11
7	13	18
6	11	31
5	12	21
4	11	13
3	9	0
2	8	0
1	2	0
Total	100	100
Number	934	934
Mean	55.	2.24
S.D.	25.	.56
S.E.	.8	.02

[a]This column divides the score range of grade-point averages in ten equal intervals from 00.0 to 4.00.

below a 1.30 grade-point average were dropped from the distribution since the comparable group had been dropped from the AFROTC sample through university regulations.

The lack of a significant difference between the mean ACE total percentile score and the mean grade-point averages for the two groups indicates that the AFROTC sample group is representative of the all male freshmen group in respect to these two factors.

The ACE total scores of the all male freshmen group and the AFROTC sample group were correlated with their grade-point averages. The resulting correlation coefficient for the sample group was found to be .420, and .400 for the all male freshmen group. The lack of a significant difference was accepted as additional evidence of the representativeness of the sample group.

II. EFFECTIVENESS OF GRADE PREDICTION

In order to determine the effectiveness with which academic success was predicted each of the three ACE scores and each of the five AFOQT aptitude composites were correlated with grade-point averages.

ACE composites versus grade-point averages. The ACE quantitative, verbal, and total scores were correlated with grade-point averages with results as shown in Table IV. The correlation coefficients between the ACE scores and grade-point averages of the sample group were: quantitative, .445; linguistic, .340; and total, .420. Contrary to expectancy the quantitative score had a higher coefficient of correlation with

TABLE IV

CORRELATIONS OF THE ACE QUANTITATIVE, LINGUISTIC, AND TOTAL PERCENTILE SCORES WITH GRADE-POINT AVERAGES

(1) Composite	(2) N	(3) Coefficient of correlation	(4) Standard error
Quantitative	140	.445	.067
Linguistic	140	.340	.074
Total	142	.420	.069

accumulative grade-point averages than did the linguistic score, however, the difference was not statistically significant.

There is some evidence to suggest that the higher predictive validity of the quantitative score may be due in part to the fact that only men students were included in the study. The Stuit study of engineering students showed a correlation coefficient between quantitative scores and grades of .41 and between linguistic scores and grades of .28.[2] These results may be contrasted with those of E.E. Anderson who correlated the ACE quantitative and linguistic scores with the grade averages of 112 women students. He found a correlation coefficient between the quantitative score and grades of .39 and between the linguistic score and grades of .54.[3]

AFOQT composites versus grade-point averages. For this part of the study, scores from each of the AFOQT composites were correlated with grade-point averages. The results of these correlations are shown in Table V. Each coefficient of correlation was compared with that of the correlation of the ACE total score with grade-point averages and the statistical significance of difference determined.

---

[2] Dewey B. Stuit, et al., Predicting Success in Professional Schools (Menasha, Wisconsin: George Banta Publishing Company, 1949), p. 29.

[3] E.E. Anderson, et al., "Wilson College Studies in Psychology: I. A Comparison of the Wechsler-Bellevue, Revised Standard-Binet, and American Council on Education Tests at the College Level," Journal of Psychology, XIV (October, 1942), p. 324.

TABLE V

CORRELATIONS OF AFOQT COMPOSITE SCORES WITH GRADE-POINT AVERAGES FOR THE AFROTC SAMPLE GROUP

(1) Composite	(2) N	(3) Coefficient of correlation	(4) Standard error
Pilot aptitude	143	.212	.079
Verbal aptitude	143	.420	.068
Navigator-technical	143	.430	.068
Quantitative	143	.475	.065
Officer aptitude	143	.574	.055

Pilot aptitude. The correlation coefficient between the pilot aptitude composite and grades was .212 which was significantly below that of the ACE total score at the 1 per cent level of confidence. A relatively low correlation with grade-point averages would be expected considering the design purpose and the nature of the subtests included. Table I shows a correlation coefficient of .23 with chemistry and .14 with mathematics, both of which are quantitative subjects, but much lower correlations in the verbal areas: philosophy -.01, English .05, and history -.22. This would seem to indicate measurement of some factor common to both pilot and quantitative aptitude which is relatively independent of verbal skills. The observed correlation is much higher than that shown in Table I between the pilot score and grade averages.

Verbal aptitude. The correlation coefficient between the verbal aptitude composite and grade-point average was .420, which was equal to that of the ACE-total-score. The observed correlation compares very closely with that of the Air Force Academy study which showed a correlation coefficient of .41.

Navigator-technical. The correlation coefficient between the navigator-technical composite and grade-point average was slightly higher than that of the ACE-total-score, but not to a statistically significant degree. The observed correlation is generally in line with the average (.435) of the correlations with quantitative subjects as shown in Table I but is considerably higher than the average (.12) for correlations with verbal subjects. The above averages were computed by the writer in

accordance with the method described by Quinn McNemar.[4]

The higher correlations with quantitative subjects is an indication that the navigator-technical composite is weighted in favor of the quantitative factors which is in accordance with its design purpose.

Quantitative. The observed correlation coefficient between the quantitative composite and grade-point average was .475, which was also higher than that of the ACE-total-score, but again not to a statistically significant degree.

The quantitative composite correlated with grades at a slightly higher degree than the verbal composite but it was not to a statistically significant degree. The direction of this difference was, however, of interest. As indicated in Chapter II the ACE linguistic composite generally is expected to correlate more closely with grade averages than is the ACE quantitative composite. For this particular group the quantitative composites of both the ACE and of the AFCOT correlated more closely with grade-point averages than did the linguistic or verbal composites of either test. Table I shows that the Air Force Academy study also found quantitative scores to be more closely related to grade averages than were verbal scores. These findings may well be more indicative of the type of course work involved than the validity of the tests. If a greater part of a student's over-all course content is quantitative in nature it would be reasonable to expect the quantitative

---

[4] Quinn McNemar, Psychological Statistics (third edition; New York: John Wiley and Sons, Incorporated, 1962), p. 140.

score of either test to correlate more closely with grade-point averages.

The observed correlation coefficient between the AFOQT verbal composite and grade-point average was .420 and between the ACE linguistic composite and grade-point average, .340. The difference was not statistically significant.

The correlation coefficient observed between the AFOQT quantitative composite and grade-point average of .475 was slightly higher than that of the ACE quantitative of .445, but not to a statistically significant degree. The comparative correlations between the ACE and AFOQT quantitative and verbal/linguistic scores were tabulated and are shown in Table VI.

Officer. The correlation coefficient between the officer aptitude composite and grade-point average was .574. This was found to be significantly higher than that of the ACE-total-score at the 1 per cent level of confidence. The observed correlation was in line with the results shown in Table I which indicate a correlation coefficient of .60 between the officer composite and grade average.

TABLE VI

CORRELATIONS OF THE AFOQT QUANTITATIVE AND VERBAL COMPOSITES WITH GRADE-POINT AVERAGES IN COMPARISON WITH THOSE OF THE ACE QUANTITATIVE AND LINGUISTIC FOR THE AFROTC FRESHMEN SAMPLE GROUP

(1) Composite	(2) N	(3) Coefficient of correlation	(4) Standard error
ACE Quantitative	140	.445	.067
AFOQT Quantitative	143	.475	.065
ACE Linguistic	140	.340	.074
AFOQT Verbal	143	.420	.068

CHAPTER V

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

I. SUMMARY

This study was undertaken to determine the relative effectiveness of the AFOQT and the ACE in predicting academic success of selected students at Bowling Green State University.

In order to accomplish the stated purpose it was necessary to (1) determine the degree to which the components of both tests predicted grade-point averages; and (2) compare the relative effectiveness of prediction.

The degree to which components of both tests predicted academic success was determined by correlating the student's score for each component with his grade-point average.

In order to determine the relative effectiveness of prediction the resulting correlation coefficients of the AFOQT composites with grade-point averages were compared with those of the ACE scores and grade-point averages. The statistical significance of difference was then computed by an appropriate method.

II. CONCLUSIONS

The following conclusions were reached as a result of this study:

1. The AFOQT verbal, quantitative, and navigator-technical composites and the ACE linguistic, quantitative and total scores all

predicted academic success with relatively equal effectiveness.

The lack of a statistically significant difference between the predictive coefficients of these composites is an indication that the observed differences could reasonably be expected to have occurred as a result of chance variation and, therefore, the observed differences may not have been real.

2. The AFOQT officer composite was found to be more effective in predicting grade-point average than the ACE total score.

The evidence would seem to justify this conclusion. The AFOQT officer composite correlated with grade-point averages at a significantly higher level than did the ACE total score. The literature concerning the two tests indicated a high degree of similarity in design purpose. The fact that the AFOQT officer composite is longer and more comprehensive might well explain its higher predictive validity.

3. The relationship between the AFOQT pilot aptitude composite and the ACE total score was indeterminate.

The lack of similarity in design purpose and the significantly low correlation with grade-point averages would seem adequate to justify the stated conclusion.

4. A sufficient degree of relationship was found between the ACE and the AFOQT to warrant consideration of the ACE, or a comparable test, as an initial selection device under the proposed Air Force Officer Education Program.

This conclusion is warranted on the basis of the fact that components of both tests have shown a positive and statistically

significant relationship with grade-point averages. It is recognized that the AFOQT measures other mental and personality characteristics in addition to those measured by the ACE. However, a major design purpose of the AFOQT is to measure general intelligence or learning ability; therefore it seems reasonable to assume that any student unable to achieve an acceptable score on the ACE would also be unable to achieve an acceptable score on the AFOQT.

III. RECOMMENDATIONS FOR FURTHER STUDY

It would seem desirable to plan a more comprehensive study of the predictive effectiveness of the two tests designed to avoid the limitations of and expand the scope of this study. This might be accomplished by comparing results over a period of several years. This would increase the number of students involved and allow for broader generalizations. It would also be feasible then to correlate test scores with grade-point averages accumulated over more college work than the one semester grand-point averages possible for this study.

BIBLIOGRAPHY

SELECTED BIBLIOGRAPHY

American Council on Education. Manual of Instructions. Cooperative Test Division, Educational Testing Service. Los Angeles: 4641 Hollywood Boulevard, 1950.

American Council on Education. Psychological Examination for College Freshmen. Norms Bulletin. Los Angeles: Cooperative Test Division, Educational Testing Service, 1953.

Anderson, E.E., et al. "Wilson College Studies in Psychology: I. A Comparison of the Wechsler-Bellevue, Revised Stanford-Binet, and American Council on Education Test at the College Level," Journal of Psychology, XIV (October, 1942), 317-26.

Arnold, Frank C. "The Relationship of Selected Factors to College Success at Bowling Green State University" (Bowling Green State University, Bowling Green, Ohio, 1959). (Mimeographed.)

Berdie, Ralph F., et al. Counseling and the Use of Tests. Minneapolis: University of Minnesota, 1959.

Bloom, Benjamin S. and Frank R. Peters. The Use of Academic Prediction Scales for Counseling and Selecting College Entrants. New York: The Free Press of Glencoe, Incorporated, 1961.

Bowling Green State University. Bulletin. Bowling Green State University Bulletins, Vol. XLVII, No. 2. Bowling Green, Ohio, 1961.

Christal, Raymond E. and John D. Krumboltz. Technical Memorandum, Use of the Air Force Officer Qualifying Test in the AFROTC Selecting Program. Air Force Personnel and Training Research Center, Department of the Air Force. Lackland Air Force Base, Texas, 1957.

Crawford, Albert B. and Paul S. Burnham. Forecasting College Achievement. New Haven: Yale University Press, 1946.

Cronbach, Lee J. Essentials of Psychological Testing. Second Edition. New York: Harper and Brothers, 1960.

Fisher, R.A. Statistical Methods for Research Workers. Tenth edition. New York: Hafner Publishing Company, Incorporated, 1949.

Fleishman, Edwin A. "Psychomotor Selection Tests: Research and Application in the United States Air Force," Personnel Psychology, Vol. 9, No. 4 (Winter, 1956), 449-465.

Fowler, H.N. "American Council on Education Psychological Examination for College Freshmen," The Fifth Mental Measurements Yearbook, Oscar Krisen Buros, editor. New Jersey: The Cryphon Press, 1959.

Hagood, Margaret J. and Daniel O. Price. Statistics for Sociologists. Revised edition. New York: Henry Holt and Company, 1952.

Krumboltz, John D. and Raymond E. Christal. Predictive Validities for First-Year Criteria at the Air Force Academy. Air Force Personnel and Training Research Center, Department of the Air Force, Development Report. Lackland Air Force Base, Texas, 1957.

Lathrop, Robert L., and others. Student Attitudes Toward Reserve Officer Training Corps Programs. Bureau of Institutional Research, University of Minnesota, Minneapolis: 1962.

Lindquist, E.F. A First Course in Statistics. Revised edition. Cambridge: Houghton Mifflin Company, 1942.

MacPhail, A.H. "Q and L Scores on the ACE Psychological Examination," School and Society (September, 1942), 248-51.

McNemar, Quinn. Psychological Statistics. Third edition. New York: John Wiley and Sons, Incorporated, 1962.

Slater, Margaret M. "The Meaning of A.C.E. Scores of Students at Bowling Green State University." Unpublished Master's thesis, The Bowling Green State University, Bowling Green, Ohio, 1952.

Stuit, Dewey B., et al. Predicting Success in Professional Schools. Menasha, Wisconsin: George Banta Publishing Company, 1949.

Thurstone, L.L., and T.G. Thurstone. "The 1937 Psychological Examination for College Freshmen," Educational Record, XIX, No. 2 (April, 1938), 209-234.

Thurstone, L.L., and T.G. Thurstone. "The 1930 Psychological Examination," Educational Record, XII (April, 1931), 160-78.

Thurstone, L.L., T.G. Thurstone, and D.C. Adkins. "The 1938 Psychological Examination," Educational Record, XX (April, 1939), 263-300.

Traxler, A.C. "The Correlation Between Two Tests of Academic Aptitude," School and Society, LXI (June, 1945), 383-84.

Tupes, Ernest C., A. Carp, and Walter R. Borg. Validation of a Proposed Officer Effectiveness Selection Battery. Air Force Personnel and Training Research Center, Department of the Air Force. Lackland Air Force Base, Texas, 1957.

Tupes, Ernest C. and Raymond E. Christal. Psychological Tests and the Selection and Classification of Air Force Officers. Air Force Personnel and Training Research Center, Department of the Air Force, Development Report. Lackland Air Force Base, Texas, 1957.

United States Department of the Air Force. Air Force Officer Classification Manual 36-1. Revised. Washington: Government Printing Office, 1961.

United States Department of the Air Force. Air Force Regulation 45-48, Reserve Forces. Washington: Government Printing Office, 1958.

United States Department of the Air Force. Manual for Interpretation of Aptitude and Interest Scores of the Air Force Officer Qualifying Test. Washington: Government Printing Office, 1958.