AN ABSTRACT OF THE THESIS OF

HORTON LAWRENCE FROSS for the DOCTOR OF PHILOSOPHY
(Name) (Degree)

in EDUCATION presented on August 6, 1969
(Major) Date)

Title: AN EXPERIMENT TO ASCERTAIN THE EFFECT OF MATURATION AND EDUCATIONAL EXPERIENCE ON AIR FORCE OFFICER QUALIFYING TEST SCORES

Abstract approved: Redacted for Privacy
Franklin R. Zeran

The United States Air Force uses a test constructed by Air Force personnel to test all individuals seeking entrance into Air Force officer training programs. This test is titled "The Air Force Officer Qualifying Test" (AFOQT). There are two major classes of candidates who are required to take the test: the college graduate who applies for admittance to the Officer Training School (OTS) located at San Antonio, Texas, and the Air Force Reserve Officer Training Corps (AFROTC) cadet located in detachments on college and university campuses throughout the United States.

In the fall of 1967 officials of the Testing Branch, Headquarters, Air Force ROTC, Maxwell Air Force Base, Alabama, noticed that college graduates applying for OTS were scoring approximately twenty points higher on the AFOQT than were Air Force ROTC cadets. The question was asked, "Why should there be such a wide

variance in test scores on the same test by two seemingly comparable groups?" One answer proposed was that the college graduate was two to three years older than the cadet and in those extra years had gained maturity and valuable educational experience that enabled him to elevate his score by an average of twenty points. It was the purpose of this study to determine empirically what effect two to three years of maturation and educational experience have on scores obtained on the Air Force Officer Qualifying Test.

The test-retest method was used. Four hundred fifteen cadets who had been previously identified as either ground officer candidates or flying officer candidates and whose initial AFOQT test scores as freshmen were available to the Test Control Section, Testing Branch, Headquarters, AFROTC Maxwell Air Force Base, Alabama, were retested within 150 days of completion of degree requirements to coincide with Officer Training School eligibility. The cadets were selected from AFROTC detachments nationwide and from all type institutions to insure a valid and representative sample. After retesting, comparisons were made between Officer Quality Composite, Pilot Composite, and Navigator-Technical Composite scores and significant differences, if any, established and recorded. Twelve conclusions were made, based on the thesis data, that generally support the basic hypothesis.

An Experiment to Ascertain the Effect of Maturation and Educational Experience on Air Force Officer Qualifying Test Scores

by

Horton Lawrence Fross

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

June 1970

APPROVED:

Redacted for Privacy

Professor of Education
in charge of major

Redacted for Privacy

Dean of the School of Education

Redacted for Privacy

Dean of Graduate School

Date thesis is presented 6 August, 1969

Typed by Gwendolyn Hansen for Horton Lawrence Fross

ACKNOWLEDGEMENTS

I wish to thank the many people who helped to make this thesis possible. My deepest gratitude goes to Miss Bernice Denoise, Testing Branch, Headquarters Air Force Reserve Officer Training Command, Maxwell Air Force Base, Alabama, who shared with me the results of AFROTC Project "Round Robin." Her help was invaluable.

My deepest thanks go to Dr. Franklin R. Zeran, former Dean of the School of Education, Oregon State University, for being my major professor, guide and friend. He made it all seem like a pleasant journey.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

AN EXPERIMENT TO ASCERTAIN THE EFFECT OF MATURATION AND EDUCATIONAL EXPERIENCE ON AIR FORCE OFFICER QUALIFYING TEST SCORES

CHAPTER I

INTRODUCTION

Other than the United States Air Force Academy, the United States Air Force has both a long-term and a short-term source of potential officer candidates. The long-term source is the two year-four year Air Force Reserve Officer Training Corps programs located on approximately 173 college and university campuses. The short-term source is the college graduate who applies for admittance to the Air Force Officer Training School located near San Antonio, Texas. Both classes of candidates are required to take and successfully pass the Air Force Officer Qualifying Test (short title: AFOQT). The test is administered to the Air Force Reserve Training Corps (short title: AFROTC) cadets in their freshman or sophomore years (by headquarters, AFROTC directive) and to college graduates who seek entrance into Officer Training School (short title: OTS).

Statement of the Problem

In the fall of 1967, officials at the Testing Branch, Headquarters, Air Force ROTC, Maxwell Air Force Base, Alabama, noticed that

college graduates applying for OTS were scoring approximately twenty points higher on the AFOQT than were Air Force ROTC cadets.

Purpose of the Study

The question was asked, "Why should there be such a great difference in test scores on the same test by two apparently comparable groups?" One answer proposed was that the college graduate was two to three years older than the cadet and in those extra years had gained maturity and valuable educational experience that was enabling him to elevate his score by an average of twenty points. It was the purpose of this study to determine empirically what effect several years of maturation and educational experience have on scores obtained on the AFOQT.

Definitions

AFOQT = Air Force Officer Qualifying Test

AFROTC = Air Force Reserve Officer Training Corps

Experience = Participation in events

GATB = General Aptitude Test Battery, U. S. Employment Office Test

Maturation = A vaguely defined process which may refer to (1) practical wisdom (intellectual maturation) in contrast to intelligence, (2) steady and socially acceptable emotional behavior (emotional

maturation) or (3) mastery of effective social technique (social maturation). The term may be relative to chronological age; the arbitrarily set period between 21 and 65 years.

OTS = Officer Training School for Air Force Officer candidates, San Antonio, Texas

Test - retest = Administering the same test to the same group again.

CHAPTER II

REVIEW OF THE LITERATURE

The Air Force Officer Qualifying Test

Very little information concerning the Air Force Officer Qualifying Test has been made available to the general public, since it is a product of Air Force research and is used solely by the Air Force to test Air Force personnel and persons seeking entry into Air Force officer programs. Background information concerning the AFOQT is contained in a 35-page report made by Mr. Raymond E. Christal and John D. Krumboltz, 1st Lt. USAF, dated February, 1957 and prepared at the Personnel Laboratory, Air Force Personnel and Training Research Center, Lackland Air Force Base, Texas. This report is not available to the public but may be found in some AFROTC detachment historical files. This report states (5, p. 1):

> The AFOQT grew out of the Aircrew Classification Battery (ACB) which was developed during World War II for the classification of aviation cadets into the various aircrew specialties. The aircrew battery proved to have high validity for predicting the success of cadets in pilot training as can be seen in Figure 1. After the war, there were a number of changes made in the pilot training program. However, the Aircrew Classification Battery was periodically revised and maintained its predictive validity (see Figure 2).
>
> When the AFROTC program was reoriented to become a major source for Air Force pilot and observer

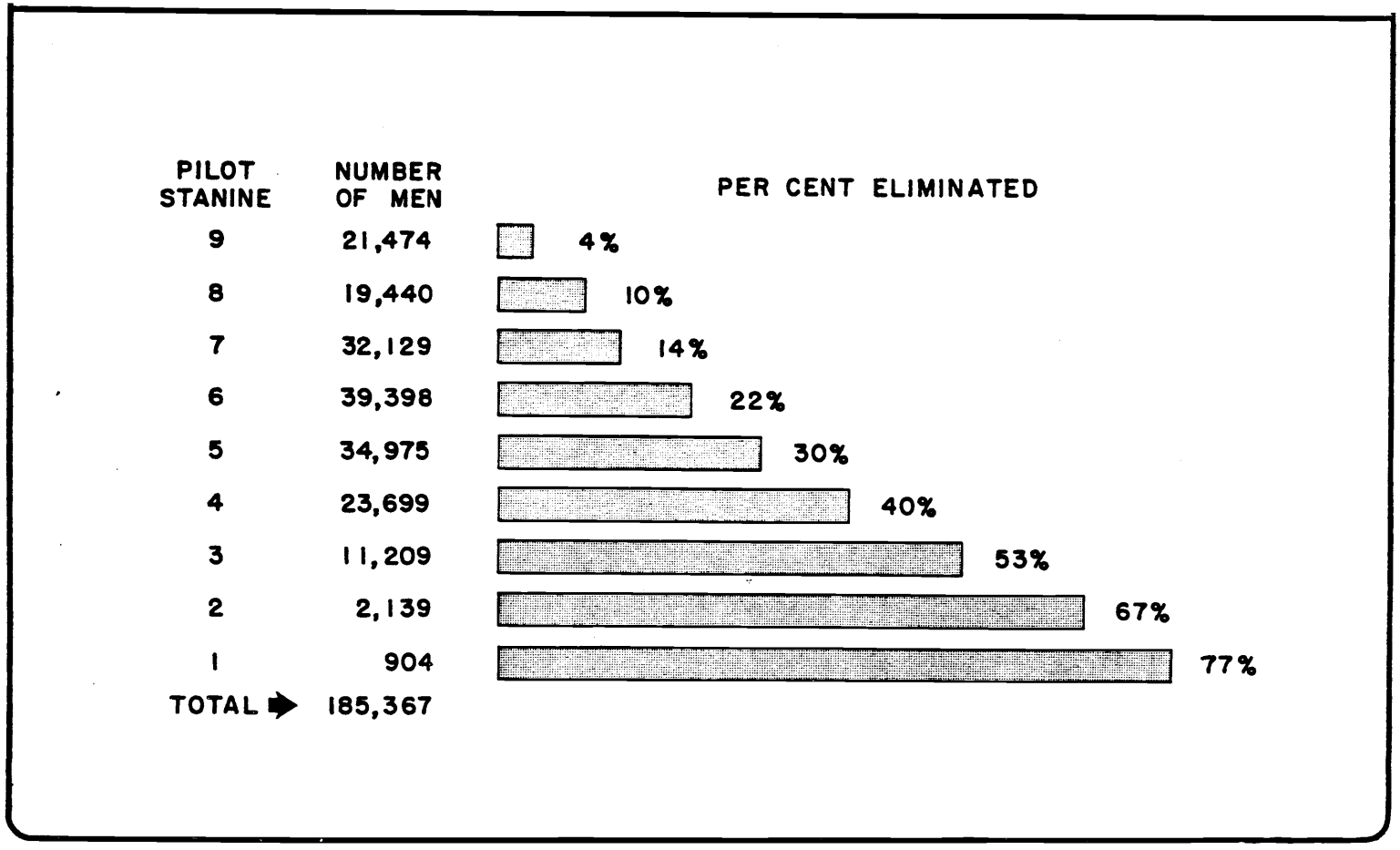

Figure 1. Effectiveness of Pilot Stanine in Predicting Elimination of Aviation Cadets in Primary Pilot Training During World War II.

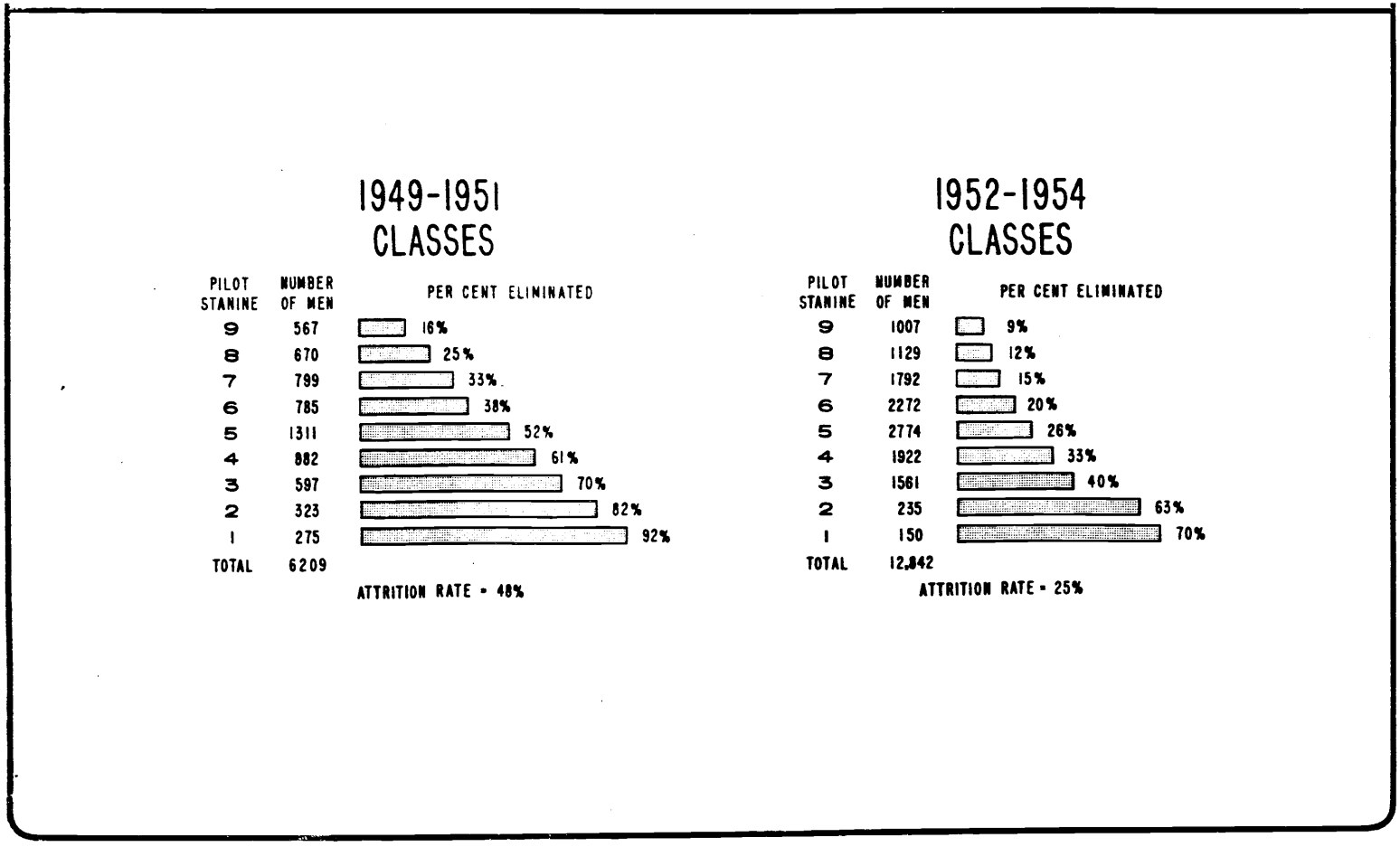

Figure 2. Effectiveness of Pilot Stanine in Predicting Elimination of Aviation Cadets in Primary Pilot Training.

personnel, the question arose as to whether cadets in this program should be selected for advanced standing on the basis of scores from a test similar to the aircrew battery. Some individuals felt that testing would be unnecessary because any college graduate should have sufficient talent to complete successfully the pilot or observer training programs. Besides, it had been a long-standing policy to accept officers on active duty into flying-training programs without aptitude testing. However, there was ample evidence that the Pilot Stanine was as valid for college graduates as for less highly educated populations. From the information presented in Figure 3, it can be seen that even within a group as highly selected and trained as West Point graduates, the Pilot Stanine has substantial validity. Although this group tends to score higher than other groups on the Pilot Stanine, those few who did receive low stanine scores had a high elimination rate in training. Also, several large-scale experimental studies had indicated that student officers with low Pilot Stanines had very little chance of successfully completing the pilot training program. On the basis of this information, a decision was made to use selection tests in the AFROTC program.

The first form of the AFOQT (Form A) was a simplified version of the Aircrew Classification Battery. Unlike the ACB, it did not contain the well-known psychomotor (apparatus) tests. Although these tests would have added validity to the paper-and-pencil tests, they would have been too expensive to administer at the 200 AFROTC detachments which were in operation in 1953. A second difference was the omission of any speeded subtests. This, too, was expected to result in the loss of some validity, but this outcome had to be weighed against the possibility of serious timing errors in the administration of highly-speeded tests by relatively untrained examiners.

The Personnel Laboratory of 1957 is now the 6570th Personnel Research Laboratory, Aerospace Medical Division, Air Force Systems Command, Lackland Air Force Base, Texas. This laboratory is responsible for the construction and maintenance of the AFOQT.

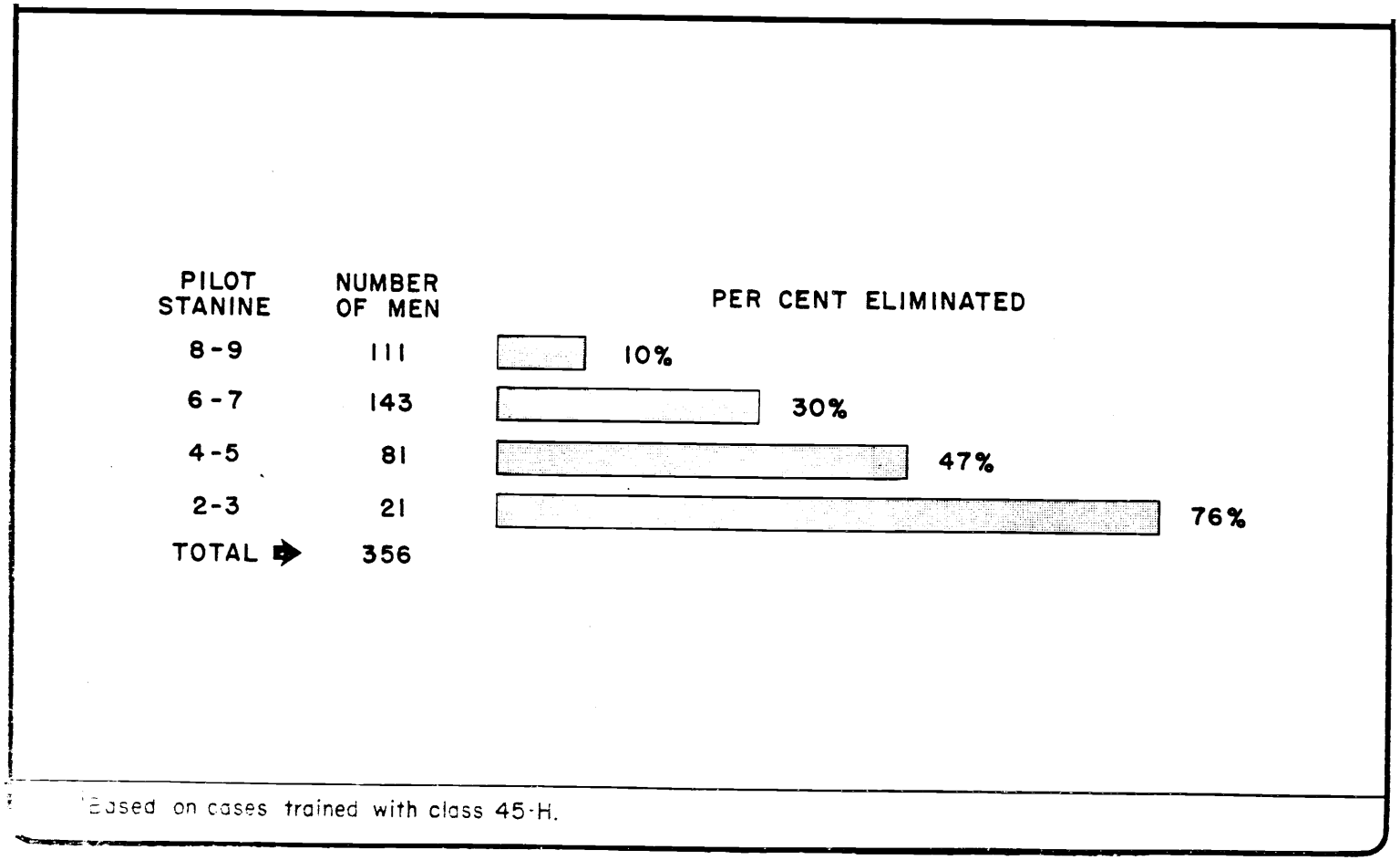

Figure 3. Effectiveness of Pilot Stanine in Predicting Elimination of West Point Cadets in Primary Pilot Training.

GATB Longitudinal Maturation Study

A review of the literature revealed a longitudinal study of maturation made by Robert C. Droege (8), U. S. Employment Service, Branch of Occupational Test Development, Washington, D. C. Droege states that the U. S. Employment Office had conducted three large-scale longitudinal studies to increase the usefulness of the General Aptitude Test Battery (GATB). His report covers the first in the series of tests. The desire was to increase the usefulness of the GATB for counseling high school students. He reports that maturation increases were largest between the ninth and twelfth grade and smallest between the eleventh and twelfth grade.

CHAPTER III

METHODOLOGY

The test-retest method was used. Approximately 400 cadets who had been previously identified as either ground officer candidates or flying officer candidates and whose initial AFOQT test scores as freshmen were available to the Test Control Section, Testing Branch, Headquarters, AFROTC Maxwell Air Force Base, Alabama, were retested within 150 days of completion of degree requirements to coincide with Officer Training School eligibility. The cadets were selected from AFROTC detachments nationwide and from all type institutions to insure a valid and representative sample. After retesting, comparisons were made between Officer Quality Composite, Pilot Composite, and Navigator-Technical Composite scores and significant differences, if any, established and recorded.

CHAPTER IV

PROCEDURES

The Test

The AFOQT is composed of several "booklets." Test scores obtained on these "booklets" may be combined in several ways to obtain "composite" scores. Only the Officer Quality Composite, the Pilot Aptitude Composite and the Navigator/Technical Composite were of concern to this study. Details concerning the structure and standardization of the test as well as its many statistical characteristics are available in a variety of publications of the 6570th Personnel Research Laboratory, Aerospace Medical Division, Air Force Systems Command, Lackland Air Force Base, Texas, which organization is responsible for the construction and maintenance of the test.

The Sample

One hundred seventy three collegiate institutions host Air Force ROTC. Annually, more than 40,000 cadets are administered the AFOQT on these campuses. For the purpose of this study a representative sample composed of 32 of these institutions was chosen. Factors considered were geographic spread, general aptitudinal level as demonstrated by past AFOQT scores,

compulsory/elective programs, size of institution, and type of institution (land grant, private, church affiliated, municipal, military, etc.).

In February 1968 the Air Science 400 senior class rosters at these 32 institutions were screened. Records revealed that AFOQT scores in all three composites were available for more than 400 of these senior cadets upon their initial testing, either as freshmen or as sophomores. It was decided to re-administer the AFOQT to as many of this group as possible when they were within five months of graduation and to use their obtained scores at the two administrations for the purpose of this study.

The Test Forms

At the time that the members of the experimental group initially took the AFOQT, the 1964 version (AFOQT-64) was in use. The test used in the final administration (in the spring of 1968) was the 1966 version (AFOQT-66). The 6570th Personnel Research Lab in its literature has repeatedly given assurances that successive forms of the AFOQT are identical in terms of psychological functions measured. Differences between successive forms are largely a matter of item content and form and minor administrative adjustments. Representatives of the 6570th Personnel Research Lab have stated that the scores obtained on AFOQT-64 and AFOQT-66 are essentially

comparable and that scores achieved on the two retain a consistency of meaning. Further, successive forms of the AFOQT are equated in terms of item statistics. This comparability between the two forms is accentuated by the fact that only percentile scores have been used in this study. In summation, were an individual to take both forms of the test at the same time, similar scores should be achieved.

Limiting Factors

It is generally conceded by educators that maturation and educational experience should have an elevating effect on the score achieved by an individual on the AFOQT. The critical point seems to be in the answer to the question, "How Much?". In the comparisons to be made in this study there are a variety of factors having a bearing on the results which will tend to dilute any actual elevation. As a consequence of the operation of these factors, statistics demonstrating overall improvement in test scores between the two administrations are sure to be very conservative. A brief discussion of three of these factors follows.

When the experimental group of cadets took the AFOQT initially (as freshmen or sophomores), they were applicants for the Advanced Course of AFROTC and a commission in the USAF. This represents an element of motivation to do well at the initial administration that would be totally absent at the final administration, for they have now

nearly completed ROTC and are on the verge of being commissioned. In addition to this lack of positive motivation at the final administration, there was a likely demotivating factor engendered by the requirement to spend the several necessary hours taking this meaningless (to them) test at a time when their minds were fully occupied with the Military Ball, the Spring Prom, final exams, graduation and initiating a career. Professors of Aerospace Studies were enjoined to counter these depressants as best they could and to motivate the seniors to do well; a difficult task.

A second factor that would surely operate to reduce demonstrated improvement in scores between the two administrations is the tendency for regression toward the mean in the upper part of the scale at the final administration. While there would be ample opportunity for improvement of test scores on the part of those who scored low on the initial administration, this opportunity diminishes rapidly for those with high initial scores. As a matter of fact, those who achieved the highest scores attainable (95th percentile) in the initial administration could not possibly improve their scores in the final administration. Any change in their scores is sure to be a negative change.

A third factor sure to reduce a demonstrable improvement in scores was the selectivity inherent in the composition of the test group. All members of this group had persisted into the senior year

of college and of AFROTC; their peers of lesser ability and aptitude had, by and large, been weeded out. Consequently, the distribution of their initial test scores should exhibit a moderate negative skew and its mean score should be significantly above the mean obtained nationally on applicants tested. Because of the associated massing tendency in the upper half of the scale, in the distribution of initial scores, a further restriction on the opportunity to increase scores has been engendered in the experiment.

Discussion of Data

By mid-June 1968, test scores obtained during the final administration were available for 415 of the seniors. It was decided to perform the analysis using these 415 sets of scores for the study. All 32 institutions selected for the sample were represented.

The 415 members of the experimental group of AS 400 cadets were distributed among cadet categories as shown in Table 1. The table also reveals the average percentage distribution, among categories, of all senior cadets over the past six fiscal years.

An examination of Table 1 reveals that selective factors have operated to cause the experimental group to be non-representative of AFROTC seniors. However, the group of institutions chosen is considered to be quite representative of the total population and statistics derived from the experimental cadet group are valid and applicable.

Table 1. Distribution, Among Categories of AS 400 Cadets, FY 63-68 nclusive, and of 415 Cadets in Experimental Sample.

Category	Seniors FY 63-68	Experimental Group	
		Number	%
IP (Pilot)	33.3	180	43.3
IN (Navigator)	6.7	38	9.2
II (Engineer, Sciences)	21.5	93	22.4
III (Other)	37.9	104	25.1

To qualify on the AFOQT for Categories II and III requires qualification only on the Officer Quality Composite. This portion of the test alone requires only 2 hours 42 minutes to administer. However, AFROTC has found that an appreciable number of cadets who initially aimed only at Category II or III experience a "change of heart" as they advance to the upper division and desire to apply for rated category. If these converts did not initially take the entire test, it is generally necessary to administer the entire AFOQT which requires 5 hours 42 minutes to administer. The entire test is necessary because part scores from successive forms are not compatible. Therefore, cadets are generally encouraged to submit to the entire test initially, whether or not they intend subsequently to apply for a rated category. It is because of this effort that scores on the entire battery were available for so many Category II and III cadets for this study.

Notwithstanding these efforts, many beginning cadets who do not plan to apply for a rated category refuse to spend the additional three hours beyond the Officer Quality Composite to complete the entire battery. In many cases to do so would require that the cadet return for the Saturday afternoon testing session. It is for this reason that the experimental group of cadets has a larger percentage in the rated categories than is customary among seniors. This does not, of course, explain why this overage in percentage has been compensated for almost entirely by a shortage in Category III.

In the graphic and tabular presentations which follow it will be noted that there is a small number of cadets whose initial test scores are lower than the minimum acceptable score for Professional Officer Course (POC) membership or their assigned category. In all these cases the individuals were readministered the AFOQT, after the required lapse of one year, and were found qualified for the POC and/or their category of enrollment. This study has employed only the scores obtained on the initial administration of the test and the final administration.

Officer Quality Composite

The Officer Quality Composite of the AFOQT consists of three parts: Quantitative Aptitude (60 items), Verbal Aptitude (60 items) and Officer Biographical Inventory (100 items). Thus it may be seen

that more than half of the test is composed of conventional scholastic aptitude test item forms. The customary distribution of Officer Quality Composite percentile scores on the part of all AFROTC applicants on a nationwide basis is customarily quite rectangular with a national mean close to 45.

Figure 4 portrays graphically the distribution of Officer Quality Composite percentile scores of the 415 members of the experimental group at the initial administration of the test. The moderate, negative skew in the distribution of these initial scores was anticipated and should be noted. Generally, elimination has been from the lower end of the scale. The mean test score of this group in the initial administration is 58.5. This is nearly 15 points higher than the annual, national mean of all applicants for AFROTC.[1] This factor, alone reduces considerably the opportunity of this group to demonstrate improvement at the final administration, by comparison with the opportunity for improvement on the part of the average AFROTC applicants.

The distribution of test scores of this same group in the final administration is seen in Figure 5. It shows striking change. Nearly one-third of the 415 members achieved the highest possible score (95)

---

[1] Source: AFROTC Testing Board, Headquarters AFROTC, Maxwell Air Force Base, Alabama.

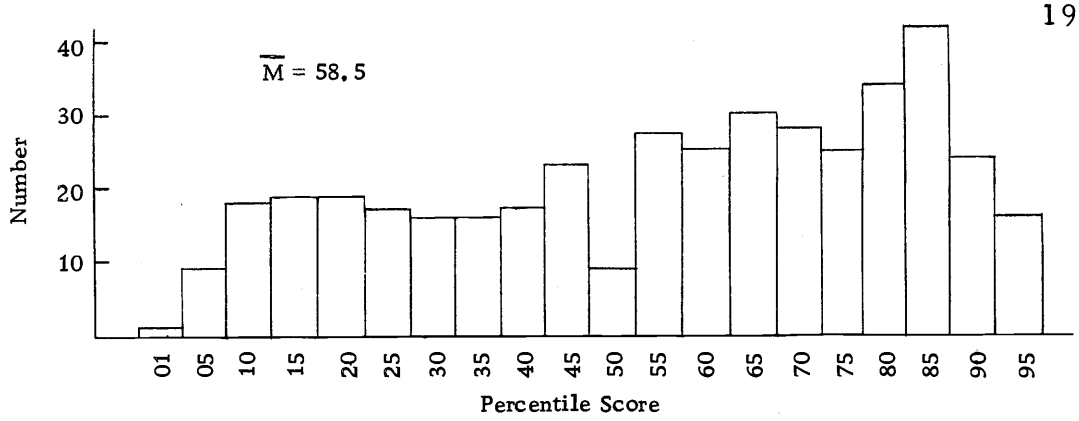

Figure 4. Distribution of Officer Quality Composite Percentile Scores--Initial Administration.

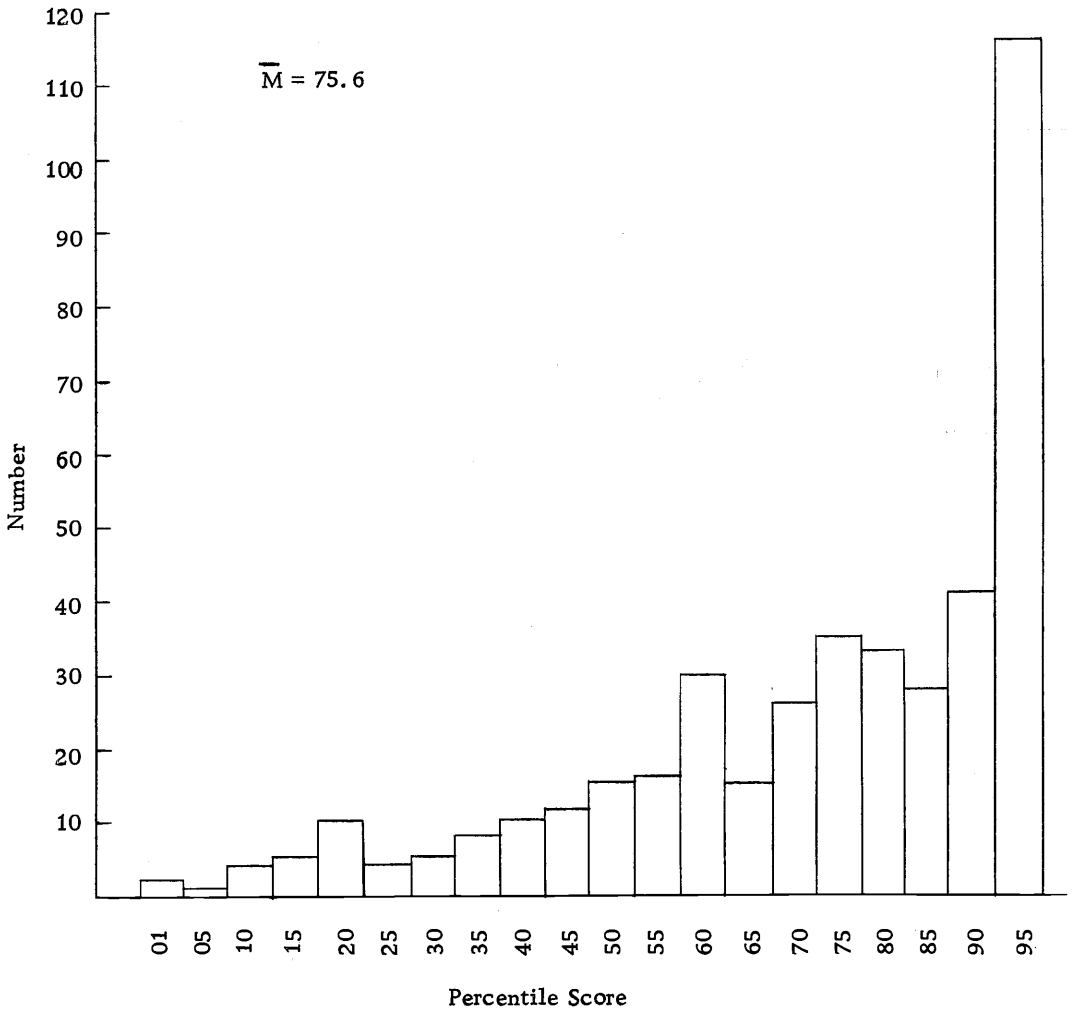

Figure 5. Distribution of Officer Quality Composite Percentile Scores--Final Administration.

on the final administration; only 16 attained this score at the initial administration. The mean score of the distribution in Figure 5 is 75.6. This is more than 30 points above the national average found in AFROTC testing, in spite of all the factors which operated to diminish measurable improvement in test scores.

Mean Officer Quality Composite percentile scores at both administrations for separate categories are shown in Table 2.

Table 2. Mean Scores by Category of Cadet on the Officer Quality Composite

	Mean Score			
Category	Initial	Final	Mean Gain	N
IP (Pilot)	58.11	76.1	18.6	180
IN (Navigator)	62.9	79.1	16.2	38
II (Engineer, Sciences)	60.0	80.9	20.9	93
III (Other)	56.2	67.7	11.5	104

There is very moderate variation in the mean scores of cadet categories at the initial administration. So little, in fact, that the categories demonstrate no significant difference in this regard. It is clear that Category III cadets' scores have not improved as much as have cadets of other categories at the final administration.

Figure 6 shows the distribution of Officer Quality Composite test score differences for the entire group. As anticipated from Figures 1 and 2, the mean test score difference is +17.1. Over 40 members

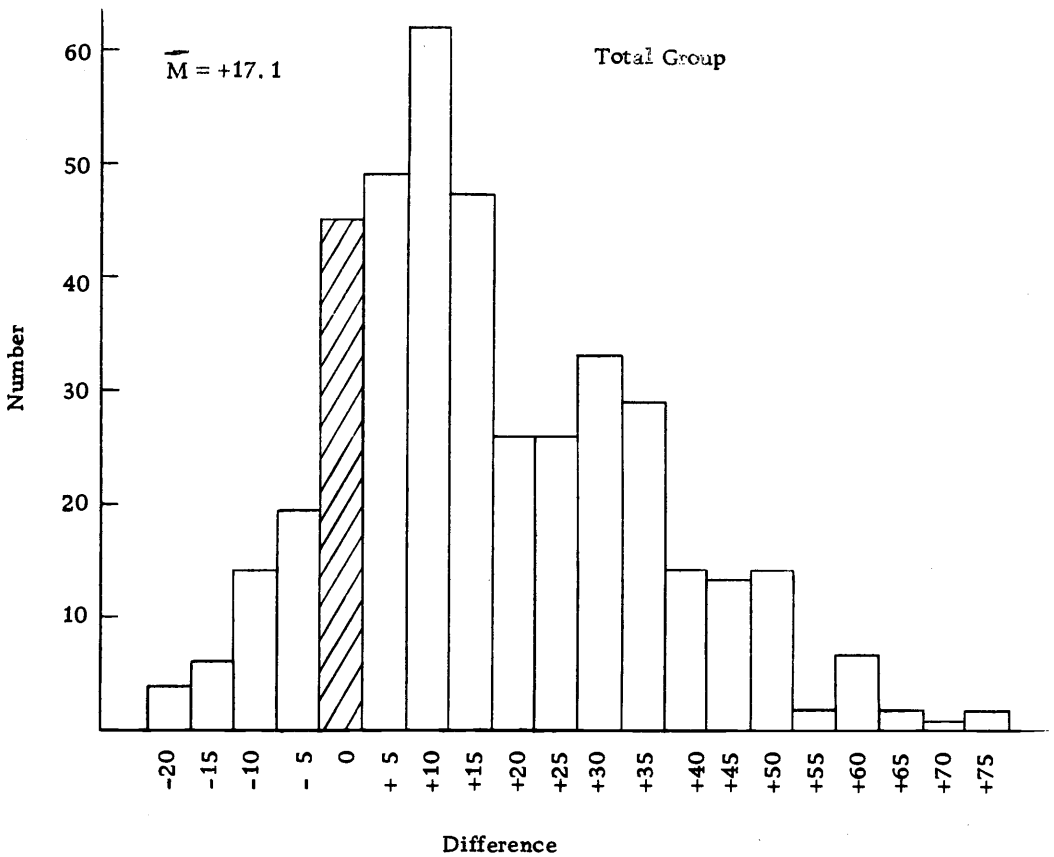

Figure 6. Difference in Officer Quality Composite Percentile Scores (Comparison between Figure 6 and 7).

of the total group had test scores in the final administration below their initial test score. It is believed that these negative differences are largely the result of the already discussed demotivating factors. It is of interest to note that 45 cadets received the same test score on both administrations.

The depressing effect of regression toward the mean is well illustrated by the inclusion of Figure 7. Here are portrayed the OQ score differences of only those whose initial scores were at or below the 50th percentile. The mean difference in this group is +28.5, nearly double that for the total group. Further evidence of regression toward the mean is to be found in a comparison of the modal scores of the two distributions. Whereas modal improvement in the total group is only 10 percentile points (Figure 6) the modal improvement in test scores for the group portrayed in Figure 7 is 35 percentile points.

Pilot Aptitude Composite

The distribution of Pilot Aptitude Composite percentile scores among the 415 members of the experimental group, at the initial administration of the test, is shown in Figure 8. Its shape is reasonably rectangular and its mean is 47.1, which compares favorably with the usual, annual mean of about 45 among applicants for the AFROTC advanced program. In other words the experimental group

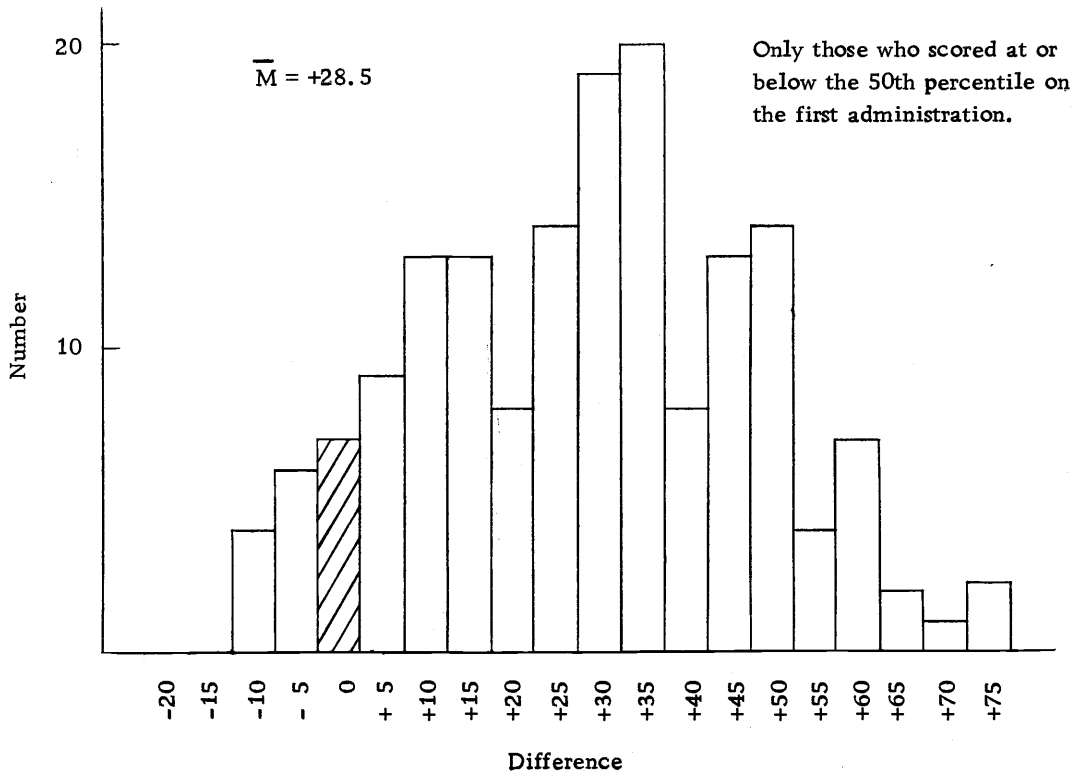

Figure 7. Difference in Officer Quality Composite Percentile Scores. (Comparison between Figure 6 and 7).

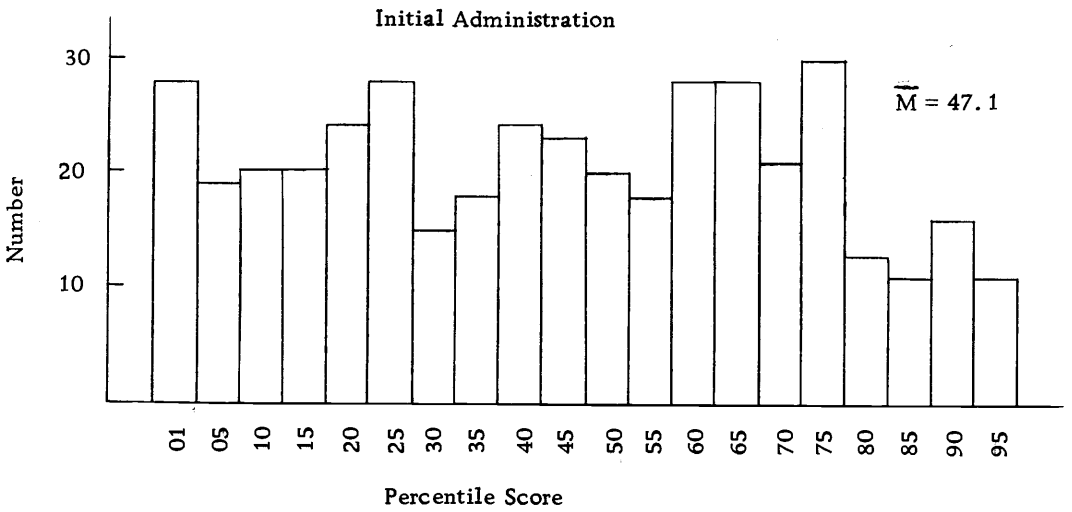

Figure 8. Distribution of Pilot Aptitude Composite Percentile Scores--Initial Administration.

demonstrates little selectivity as regards pilot aptitude and is reasonably representative of AFROTC applicants.

Figure 9 clearly demonstrates a striking improvement by the total experimental group in Pilot Aptitude Composite scores between the initial and final administrations of the test. The mean of the distribution shown in Figure 9 is 69.8, more than 20 points higher than in the initial administration.

About 25 percent of the items which compose the Pilot Aptitude Composite pertain to biographical information known to be related to success in pilot training. An additional 50 percent are directly related to knowledge of airplanes, aircraft components and operations, identification of attitude of aircraft in flight, interpretation of aircraft instrument readings and identification of maneuvers from photos of terrain. Because of their nature and the predominance of such items in the Pilot Aptitude Composite, it is patently logical to expect that those cadets who had an interest in flying would achieve significantly higher initial test scores than those without such interest even though neither group had had flying experience. To test this hypothesis Figures 10 and 11 were constructed. Although the mean in Figure 10 (41.6) is not a great deal lower than the mean in Figure 11 (53.0), the massing of scores near the zero end of the former distribution is striking.

Annually, a large number of AFROTC cadets who earnestly

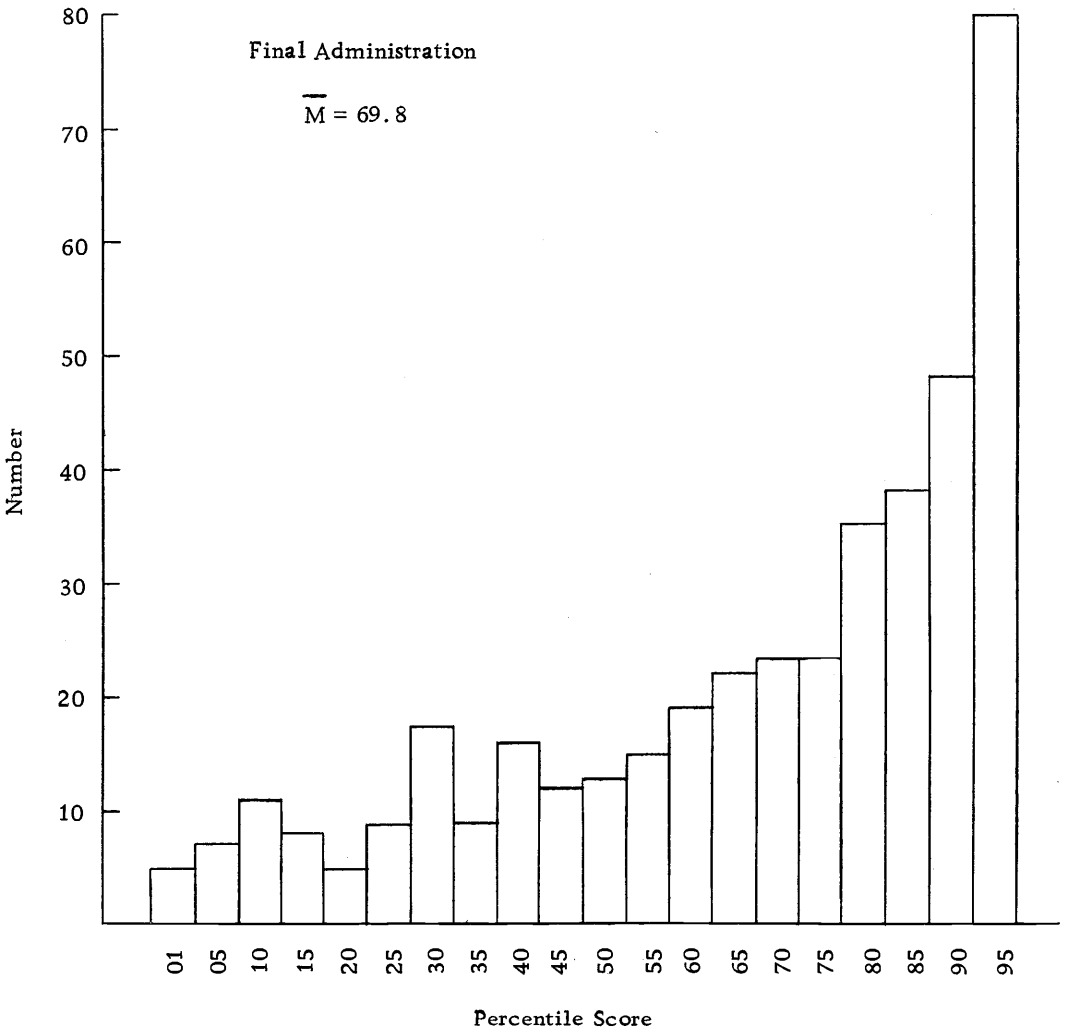

Figure 9. Distribution of Pilot Aptitude Composite Percentile Scores--Final Administration.

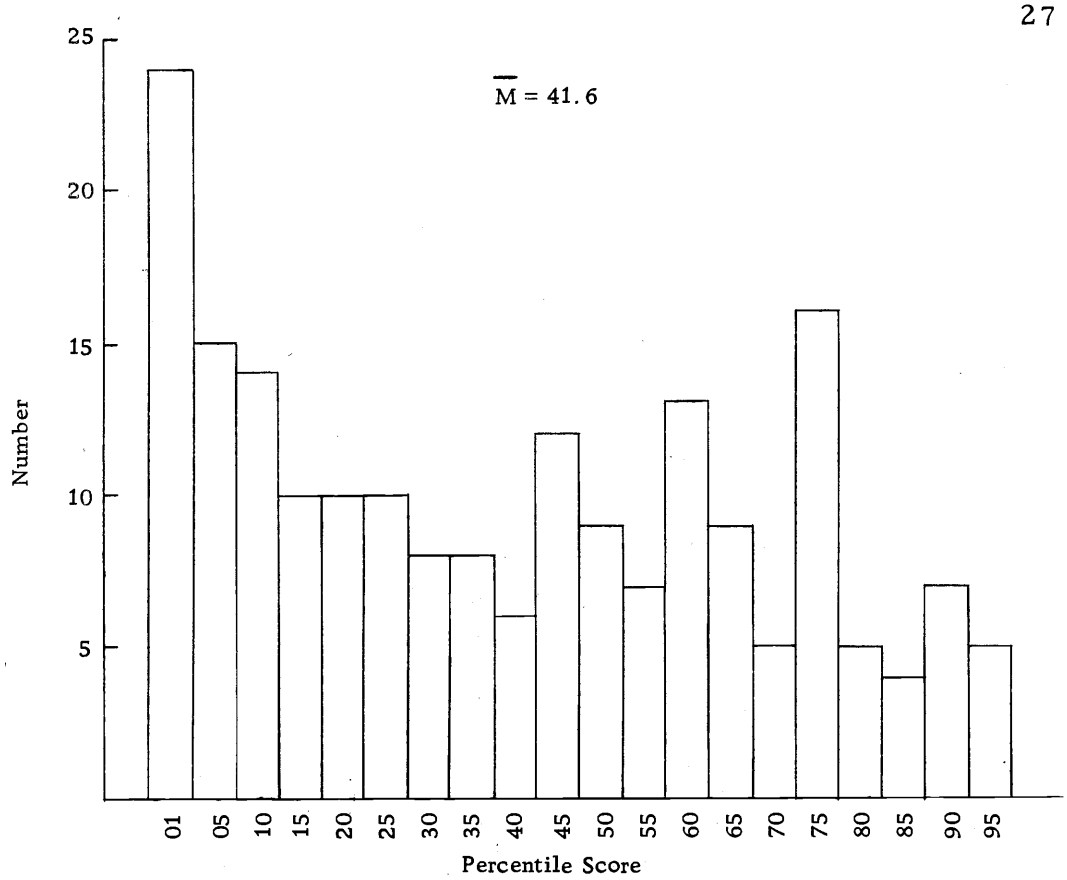

Figure 10. Distribution of Pilot Aptitude Composite Percentile Scores--Initial Administration--Category II and III Cadets Only.

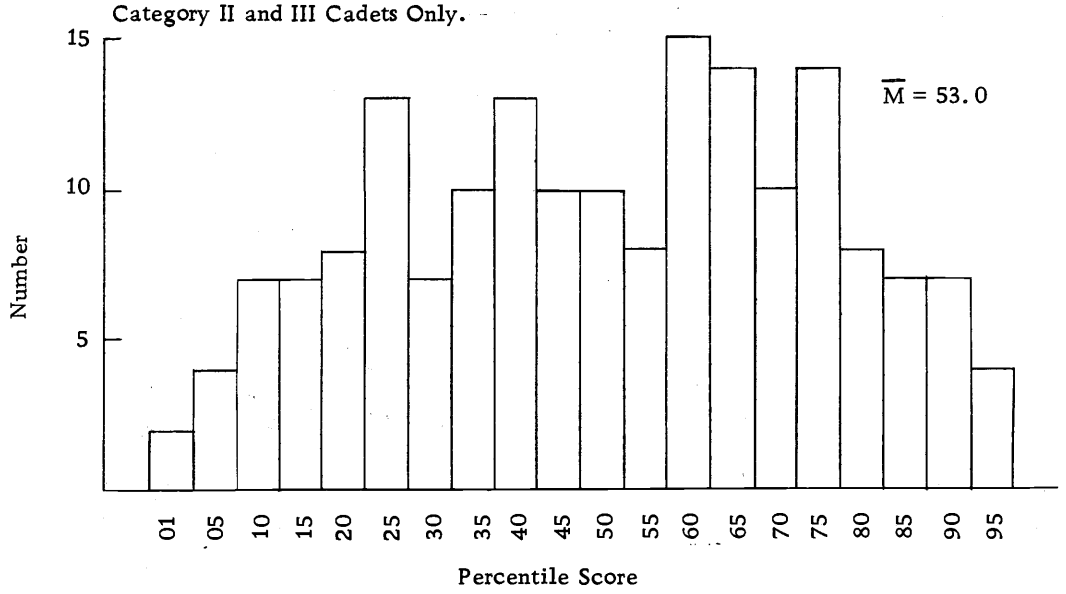

Figure 11. Distribution of Pilot Aptitude Composite Percentile Scores--Initial Administration--Category IP Cadets Only.

desire a flying category, are interested in flying and are eminently well qualified aptitudinally, are unable to meet the physical standards. Many such cadets are surely included in the data presented in Figure 11. One can speculate that they are largely to be found in the upper half of the distribution and that without them the distribution would have a lower mean and a much more marked skew.

When the Pilot Aptitude Composite percentile scores obtained in the final administration are separated into the two groups (Categories II and III on the one hand and Category IP cadets on the other), Figures 12 and 13 show the result. Quite understandably the two groups are now shown to be considerably more disparate. The mean score in Figure 12 is 59.4, in Figure 13 it is 80.9. It is quite apparent that improvement in test scores for many Category IP cadets has been restricted by their inability to attain a score higher than that represented by the 95th percentile. Nearly 30 percent of Category IP cadets achieved this maximum possible score in the final administration. Two-thirds of the group were at or above the mean score.

A comparison of Figures 10 and 12 shows that Category II and III cadets have improved their Pilot Aptitude Composite percentile scores by an average of almost 20 points. It is apparent that the distribution of test scores has shifted from a high positive skew to a significant negative skew. Whereas 24 of these cadets scored below the 5th percentile and only 5 scored at or about the 95th percentile at

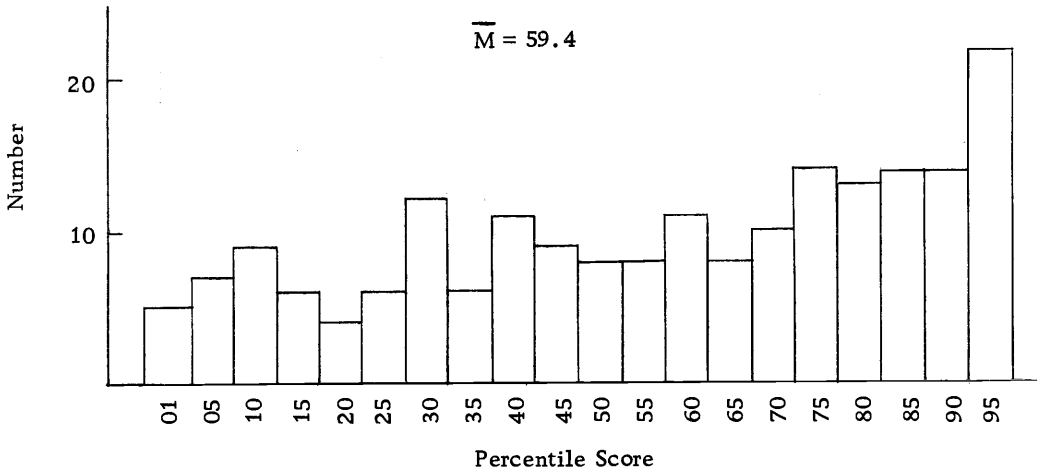

Figure 12. Distribution of Pilot Aptitude Composite Percentile Scores--Final Administration--Category II and III Cadets Only.

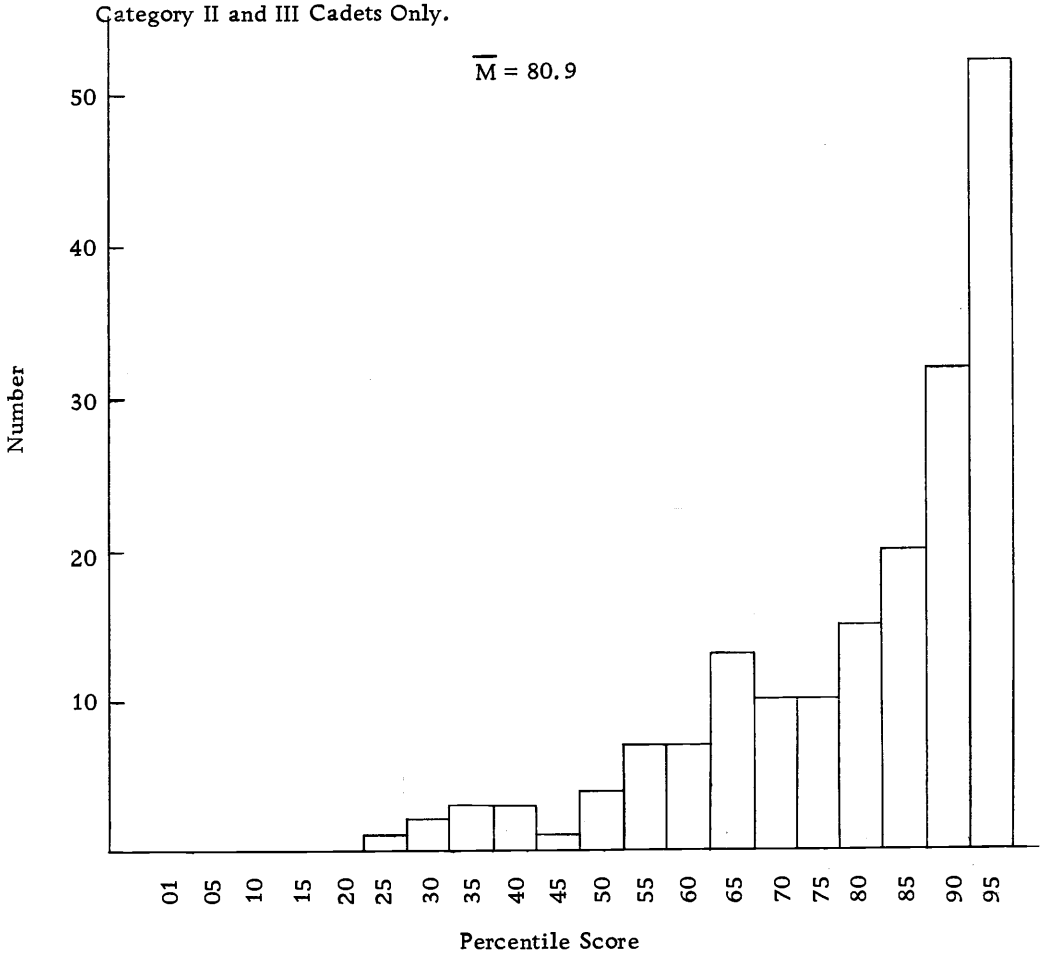

Figure 13. Distribution of Pilot Aptitude Composite Percentile Scores--Final Administration--Category IP Cadets Only.

the initial administration, these numbers have virtually been reversed in the final administration. These results seem surprising when one considers that 75 percent of the items which compose the Pilot Aptitude Composite are directly related to flying.

The remaining 25 percent of the items in this composite test knowledge of mechanical principles and are closely related to work in many engineering fields. Since many of the Category II cadets in the experimental group are in colleges of engineering, it was thought well to test whether their educational experiences were making a substantial contribution to the improvement in Pilot Aptitude Composite scores among non-flying cadets. To do this the 197 members of this group were separated into their respective categories and Figures 14 and 15 were constructed. A comparison of the two distributions and of their means fails to support any significant difference between the groups in terms of score improvement in the Pilot Aptitude Composite.

It is quite possible that this significant improvement in test scores on the part of these non-flying cadets is the result of instruction received during Aerospace Studies courses, a portion of which has to do with aerodynamics and principles of flight. A comparison of Pilot Aptitude Composite test scores as freshmen and again as seniors in a large representative group of non-Air Force ROTC cadets should prove interesting.

A demonstration of the depressing effect of regression toward

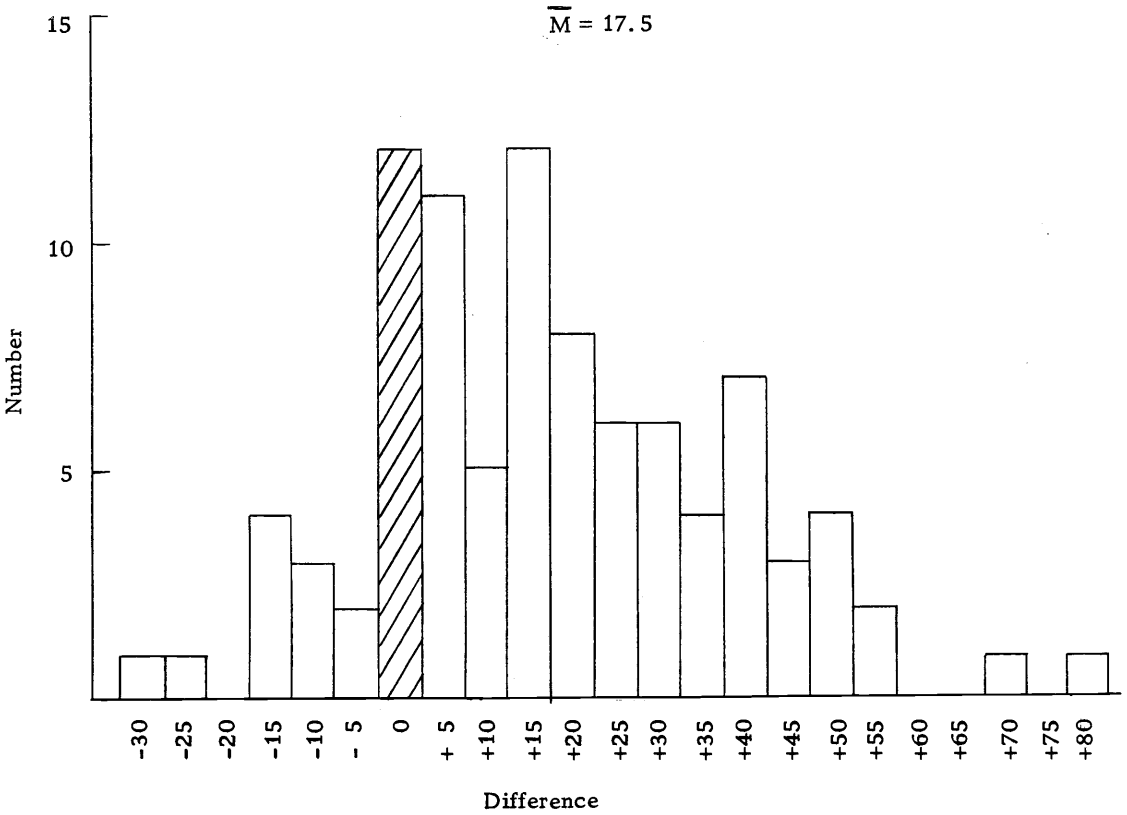

Figure 14. Difference in Pilot Aptitude Composite Percentile Scores--Category II Cadets Only.

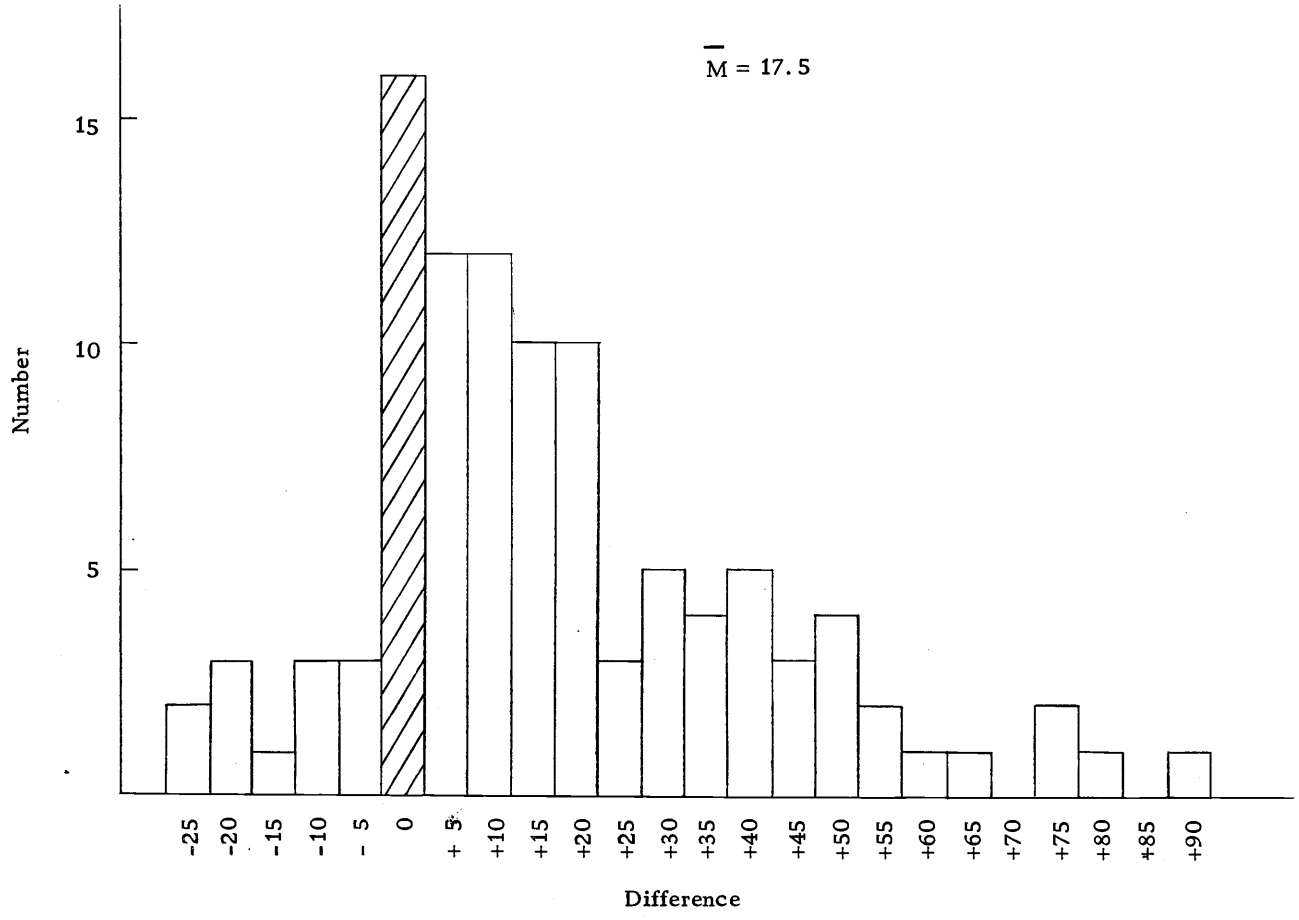

Figure 15. Difference in Pilot Aptitude Composite Percentile Scores--Category III Cadets Only.

the mean at the upper end of the scale is again furnished in Figure 16. Among the group having intervening flying training, those who initially scored below the 50th percentile in the Pilot Aptitude Composite improved their scores by an average of 46.5 percentile points in the final administration.

Navigator/Technical Composite

The Navigator/Technical Composite of the AFOQT contains 220 items. Sixty (27.3%) are common to it and to the Officer Quality Composite. Furthermore, 120 of the remaining 160 items deal with such scholastic topics as scale reading, general science, mechanical information and mechanical principles. For these reasons similarities may be expected in the distributions of scores on these two composites.

The distribution of Navigator/Technical Composite percentile scores generally found among AFROTC applicants exhibits slight positive skew with a mean customarily in the low forties. Selective factors operating through the collegiate years should produce about the same effect on the distribution of Navigator/Technical Composite percentile scores of the experimental group at the initial administration as they did on the Officer Quality Composite percentile score distribution. A comparison of Figures 4 and 17 reveals quite a bit of similarity in the distributions and almost identical mean scores.

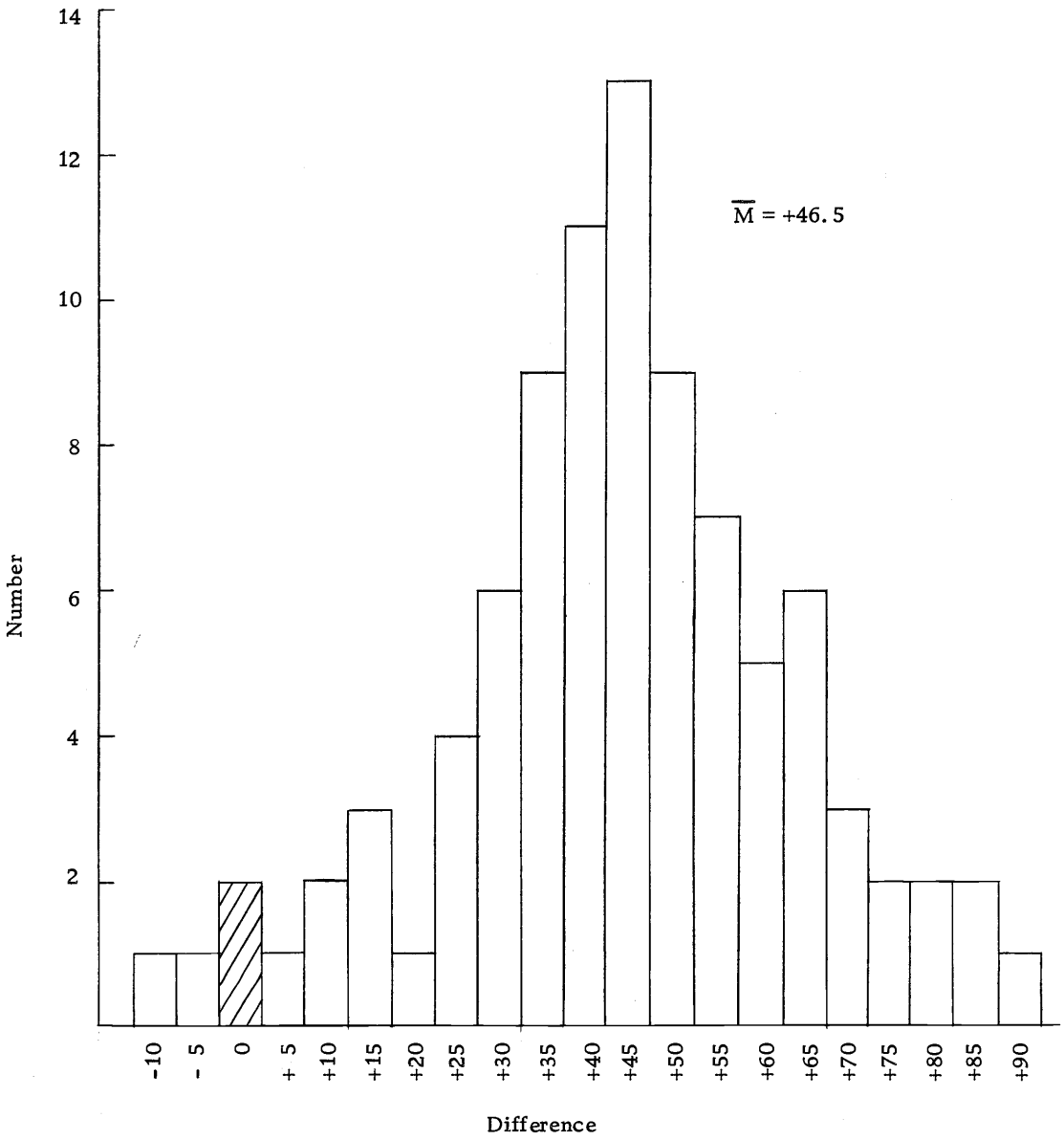

Figure 16. Difference in Pilot Aptitude Composite Percentile Scores with Intervening Flying Training--Initial Pilot Aptitude Percentile Score 50 or Less.

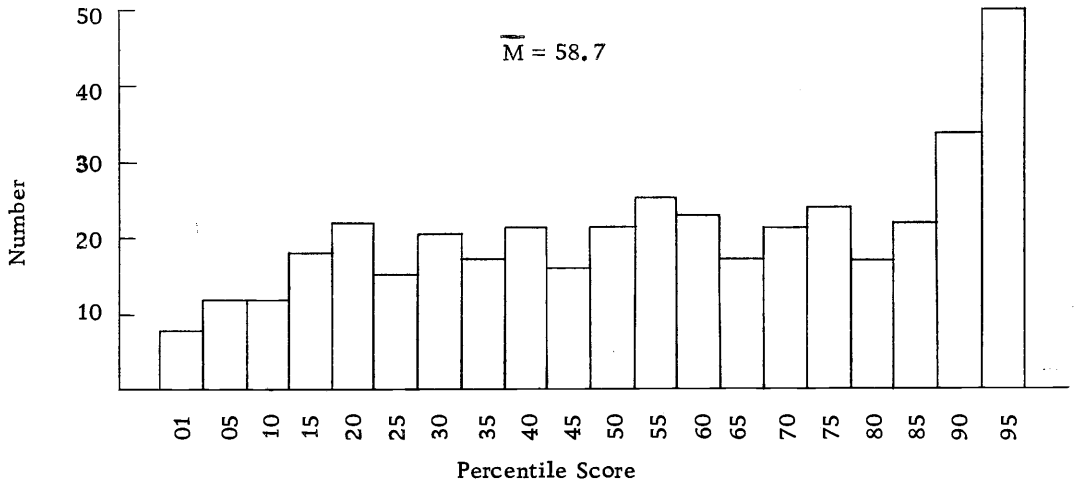

Figure 17. Distribution of Navigator/Technical Composite Percentile Scores--Initial Administration.

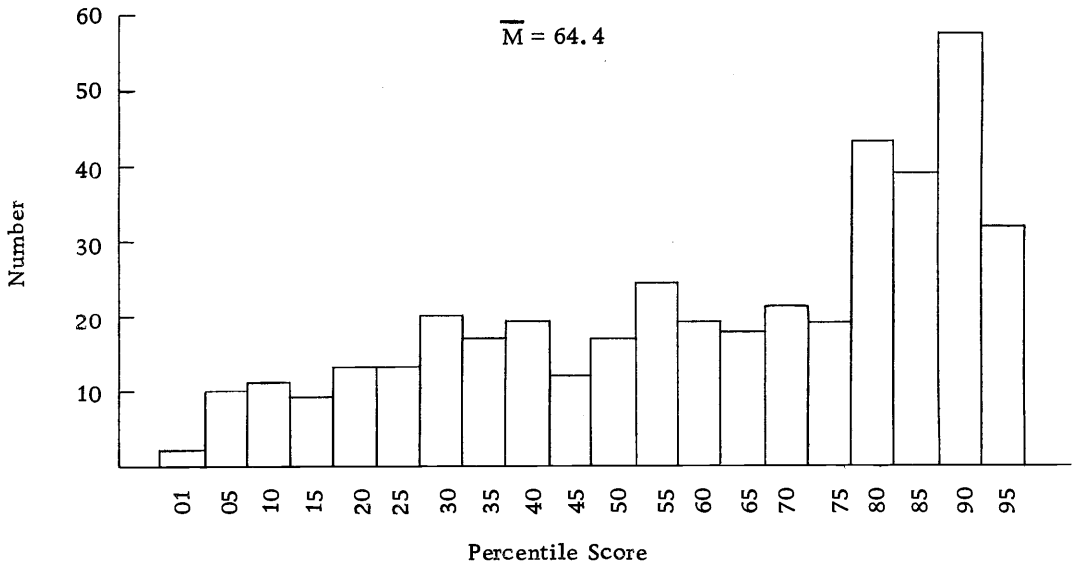

Figure 18. Distribution of Navigator/Technical Composite Percentile Scores--Final Administration.

Figure 18 reveals that the total group has demonstrated only a very modest gain (5.7 percentile points) in the mean Navigator/Technical Composite percentile score at the final administration. Of the 50 cadets who scored at the 95th percentile at the initial administration, only half (26) remained at this level in the second administration. This phenomenon alone has tended to reduce mean test score increase. Mean scores on the two administrations by cadet category are presented in Table 3.

Table 3. Mean Scores by Category of Cadet on the Navigator/Technical Composite

	Mean Score			
Category	Initial	Final	Mean Gain	N
IP	60.0	67.1	7.1	180
IN	57.8	65.4	7.6	38
II	73.2	77.0	3.8	93
III	43.8	48.0	4.2	104

The most striking features of Table 3 are the 30 percentile point differential between Category II and III cadets at both administrations and the small gain in mean score between administrations in both groups. Distribution of scores for these two categories of cadets at the two administrations are shown in Figures 19 through 22.

An examination of Figure 21 reveals that more than half the Category II cadets are at or above the 80th percentile on the initial

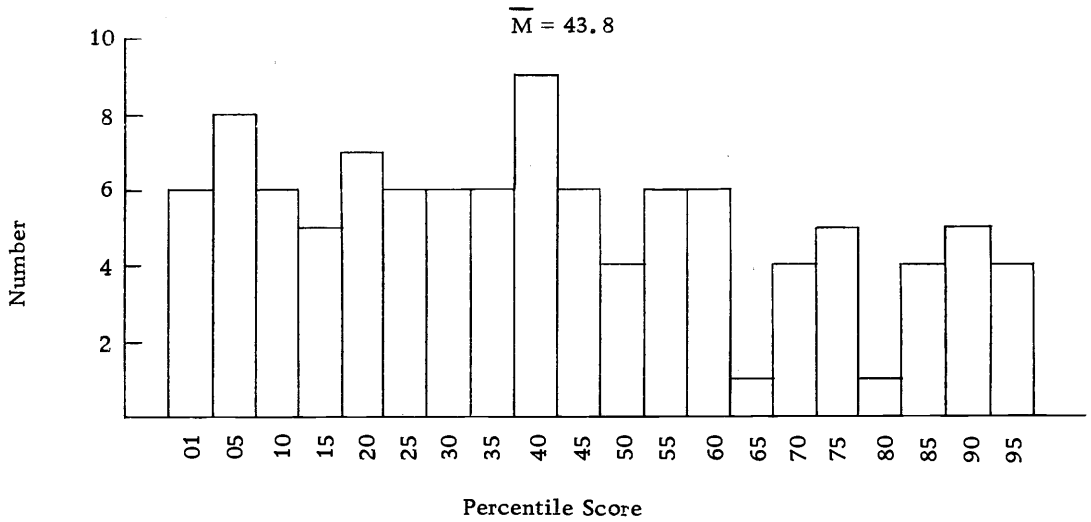

Figure 19. Distribution of Navigator/Technical Composite Percentile Scores--Category III Cadets Only--Initial Administration.

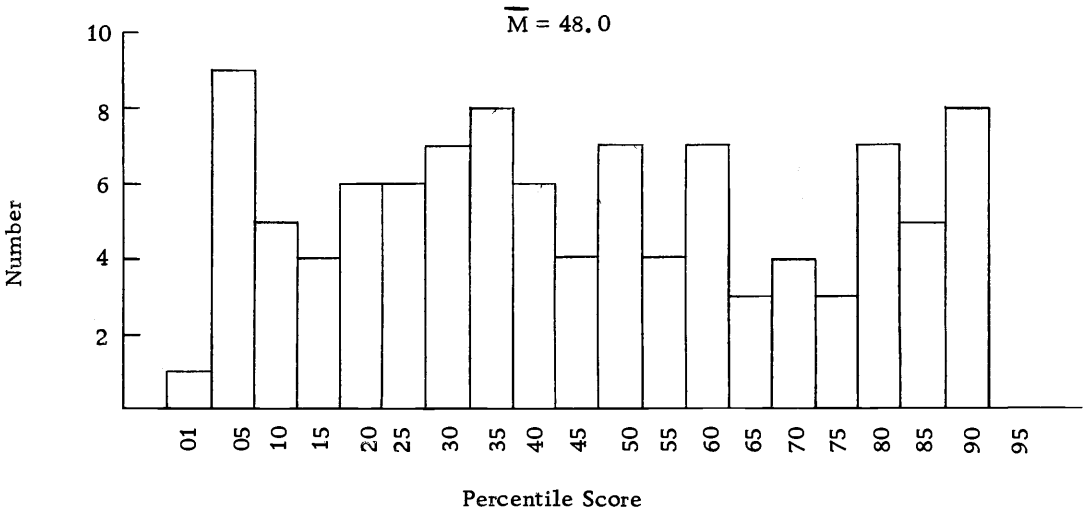

Figure 20. Distribution of Navigator/Technical Composite Percentile Scores--Category III Cadets Only--Final Administration.

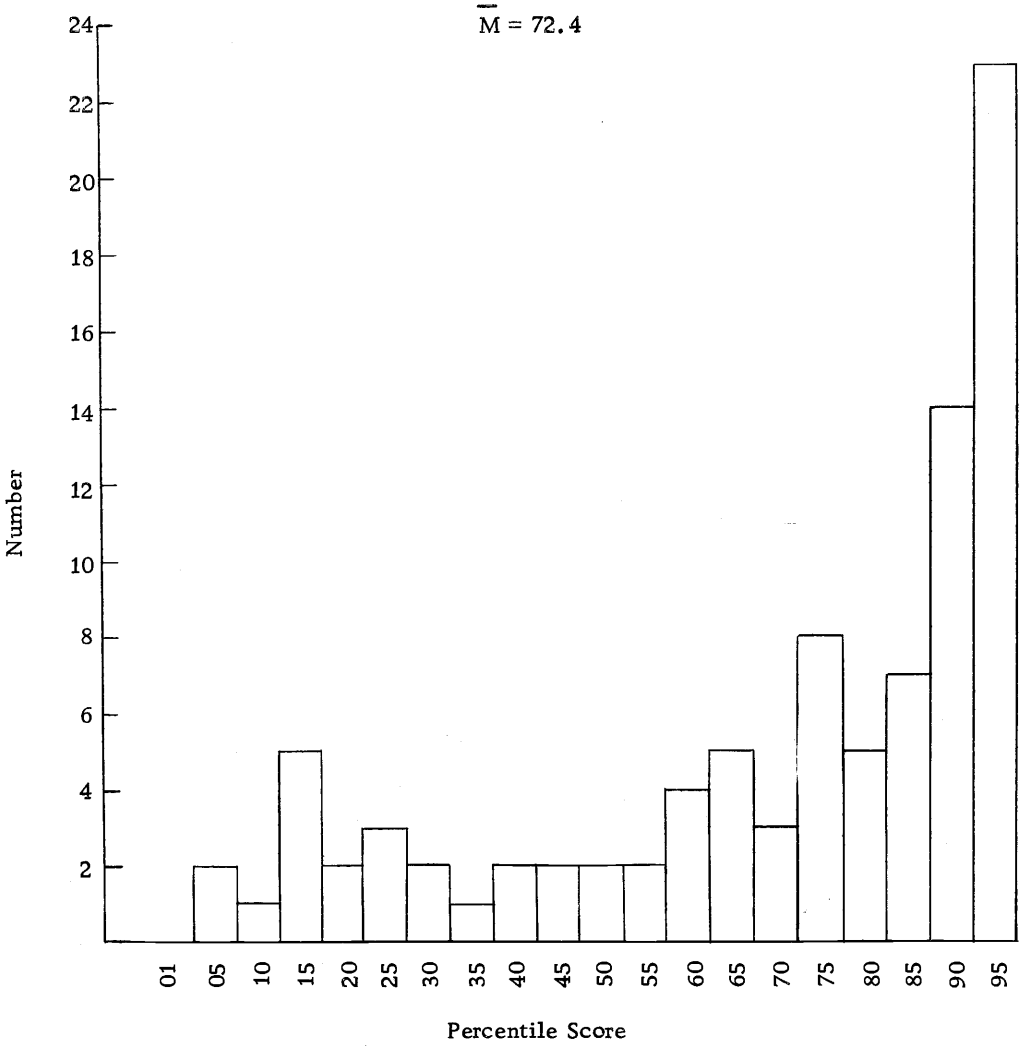

Figure 21. Distribution of Navigator/Technical Composite Percentile Scores--Category II Cadets Only--Initial Administration.

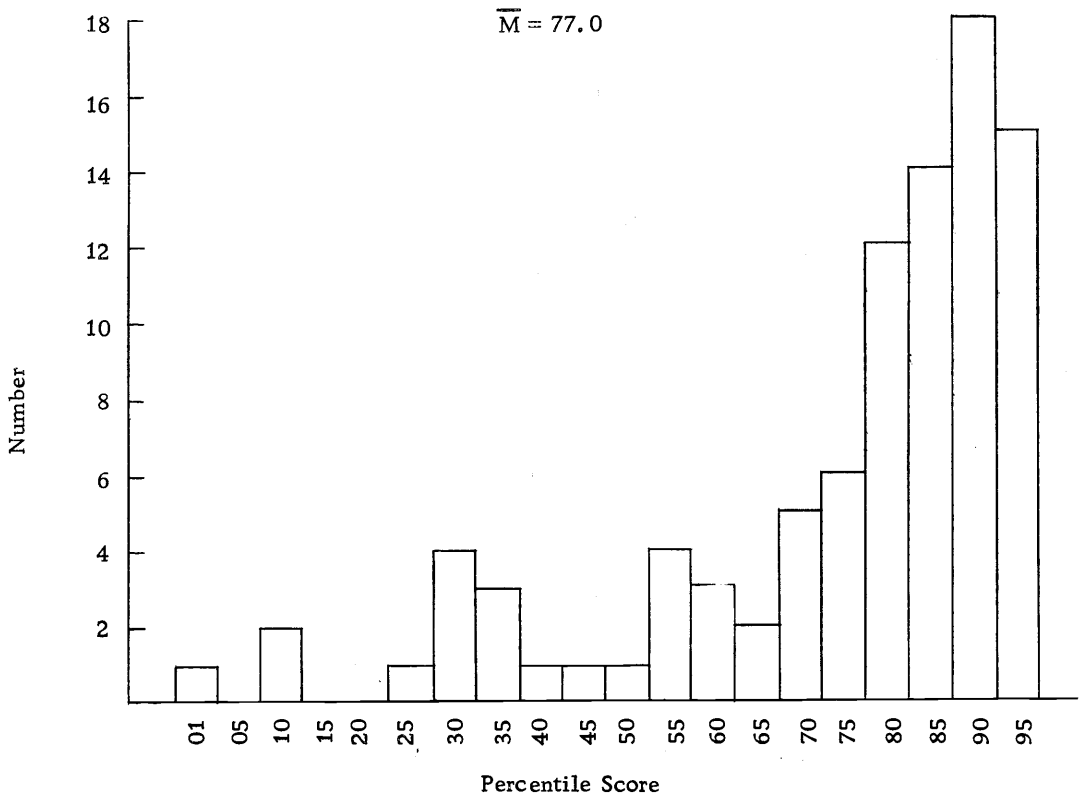

Figure 22. Distribution of Navigator/Technical Composite Percentile Scores--Category II Cadets Only--Final Administration.

test administration. It is obvious that there is little opportunity for these cadets to make significant test score improvement. A fairer comparison of the performance of Category II and Category III cadets on the two administrations of the Navigator/Technical Composite follows.

Thirty-six Category II cadets and 85 Category III cadets scored below the 75th percentile on the initial administration. Their mean scores and score gains between test administrations are given in Table 4.

Table 4. Mean Scores, at Both Administrations, of Category II and III Cadets Who Scored Below the 75th Percentile at the Initial Administration

	Mean Score			
Category	Initial	Final	Mean Gain	N
II	44.0	77.8	28.8	36
III	32.9	39.2	6.3	85

It is now apparent that when Category II and III cadets within a comparable test score range on the initial administration of the Navigator/Technical Composite are compared, Category II cadets make far greater gains at the final administration.

Extension of Study

The initial intent when this study was begun did not encompass analyses beyond those already made. However, the data provided an opportunity for certain interesting additional analyses. The extension beyond the original intent is presented below.

Institutional Differences

Because the Officer Quality Composite of the AFOQT is composed largely of items that are comparable to those found in conventional scholastic aptitude tests, it is to be expected that there would be a high correlation between Officer Quality Composite scores recorded at institutions and institutional academic requirements for admission.

Air Force ROTC has long observed that the mean Officer Quality Composite score at an institution tends to remain quite constant from year to year and that among institutions there is a wide range of means. For example, over a five-year period the mean Officer Quality Composite percentile score recorded at a certain institution ranged between 19 and 24 while at another institution during the same time period it ranged between 88 and 95. The disparity was in fact so great that the highest score recorded at the former institution during this time period was lower than the lowest score recorded at the latter.

Among the 32 institutions selected for this study, the range of mean Officer Quality Composite percentile scores in the experimental group was from 22.8 to 79.0 at the initial administration and from 42.2 to 89.5 at the final administration. Institutional means gained an average of nearly 17 points between administrations. The range of gains was from 2.1 percentile points to 35.0. Distributions of institutional means at the two administrations are shown in Figures 23 and 24.

An examination of the raw data suggested that there had been a general shifting of test score means to the right with the customary compression at the top of the scale. To test this hypothesis a correlation coefficient[2] between institutional rank-orders on the two test administrations was obtained. It was found to be .89. This high correlation coefficient demonstrates that the cadets of an institution tend to retain, over their collegiate career, about the same average rank among cadets from other institutions in the psychological traits measured by the Officer Quality Composite in spite of the diversity of educational experiences provided by the different institutions.

---

[2]When estimating the amount of correlation between institutional means, the procedure used herein employed the rank-difference method of Spearman (Rho). For finding the degree of correlation using individuals' test scores, later in this study, the Pearson product-moment coefficient of correlation has been used. The two statistics have a comparable interpretation.

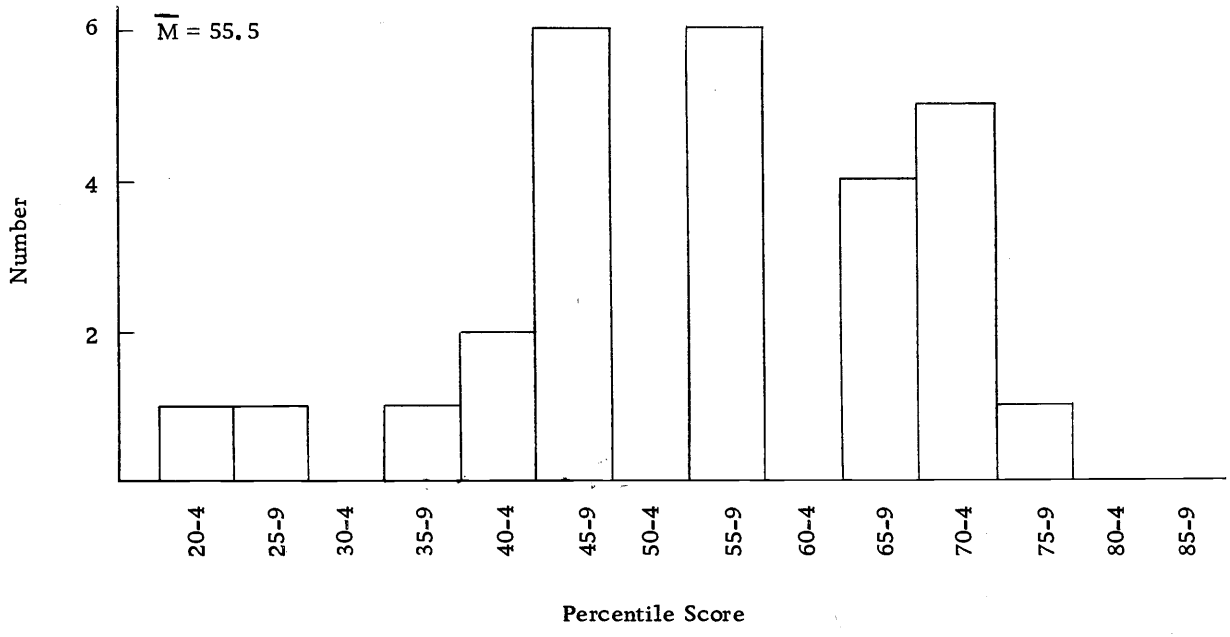

Figure 23. Distribution of Institutional Means on Officer Quality Composite--Initial Administration.

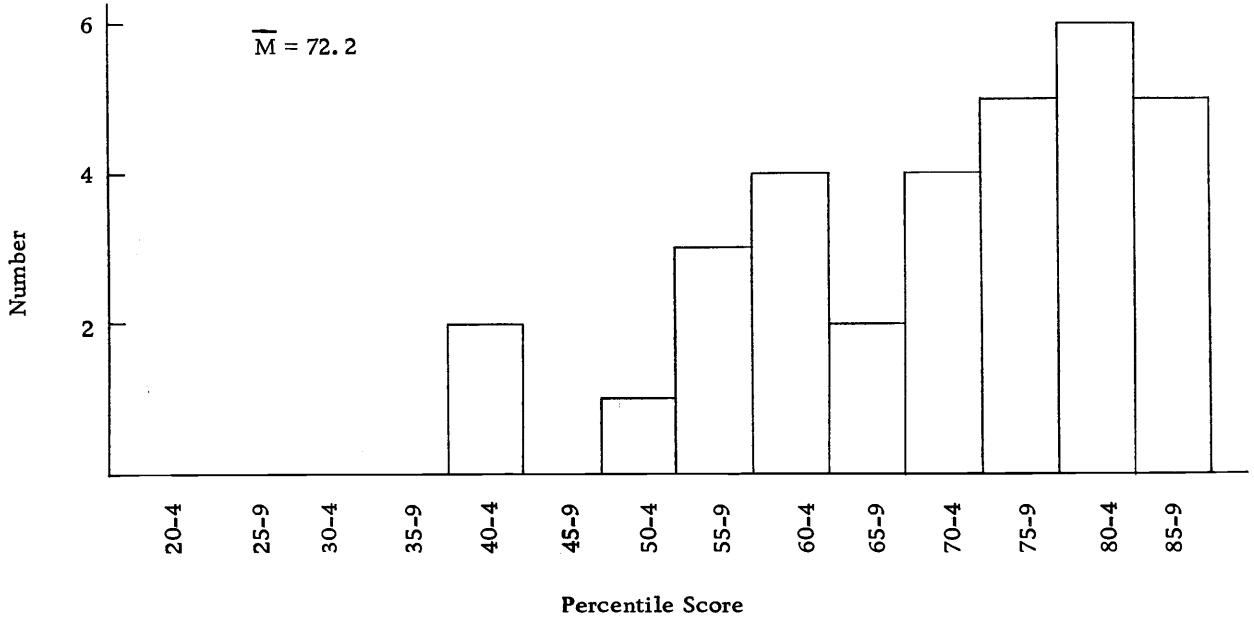

Figure 24. Distribution of Institutional Means on Officer Quality Composite--Final Administration.

The correlation coefficient between institutional means on the initial test and respective gains made between administrations was -.43. This clearly demonstrates the effect of regression toward the mean at the upper end of the scale.

The distribution of Pilot Aptitude Composite means on the two administrations of the AFOQT is shown in Figures 25 and 26. It is apparent that there is as much diversity among institutional means in the experimental group on this composite as there is on the Officer Quality Composite. The actual range of institutional means on the Pilot Aptitude Composite in the experimental group at the initial administration was from 5.5 to 74.5, which is slightly more than the range of Officer Quality Composite means, among these institutions, at the initial administration.

Figure 26 reveals that there has been considerable improvement in institutional mean scores on the Pilot Aptitude Composite at the second administration, an average gain of 23.8 points. In large measure this may be attributed to the fact that nearly half of the 415 cadets have participated in flying training between administrations. The range of means in the final administration is from 29.6 to 88.3.

To ascertain the degree to which institutions tend to retain the same relative position at the final administration as they held at the initial administration, on the Pilot Aptitude Composite, the correlation coefficient between rank orders on the two lists was obtained. It was

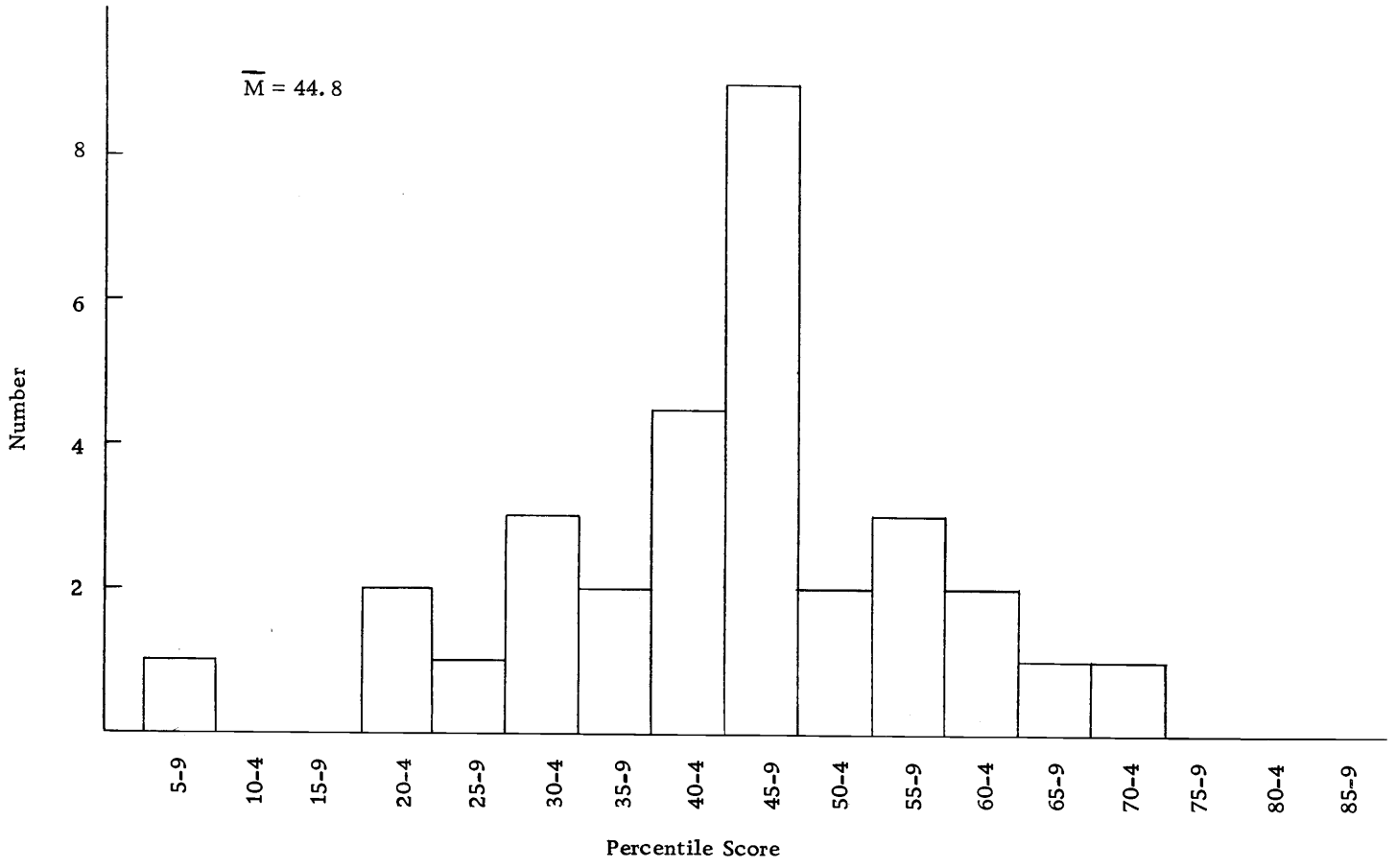

Figure 25. Distribution of Institutional Means on Pilot Aptitude Composite--Initial Administration.

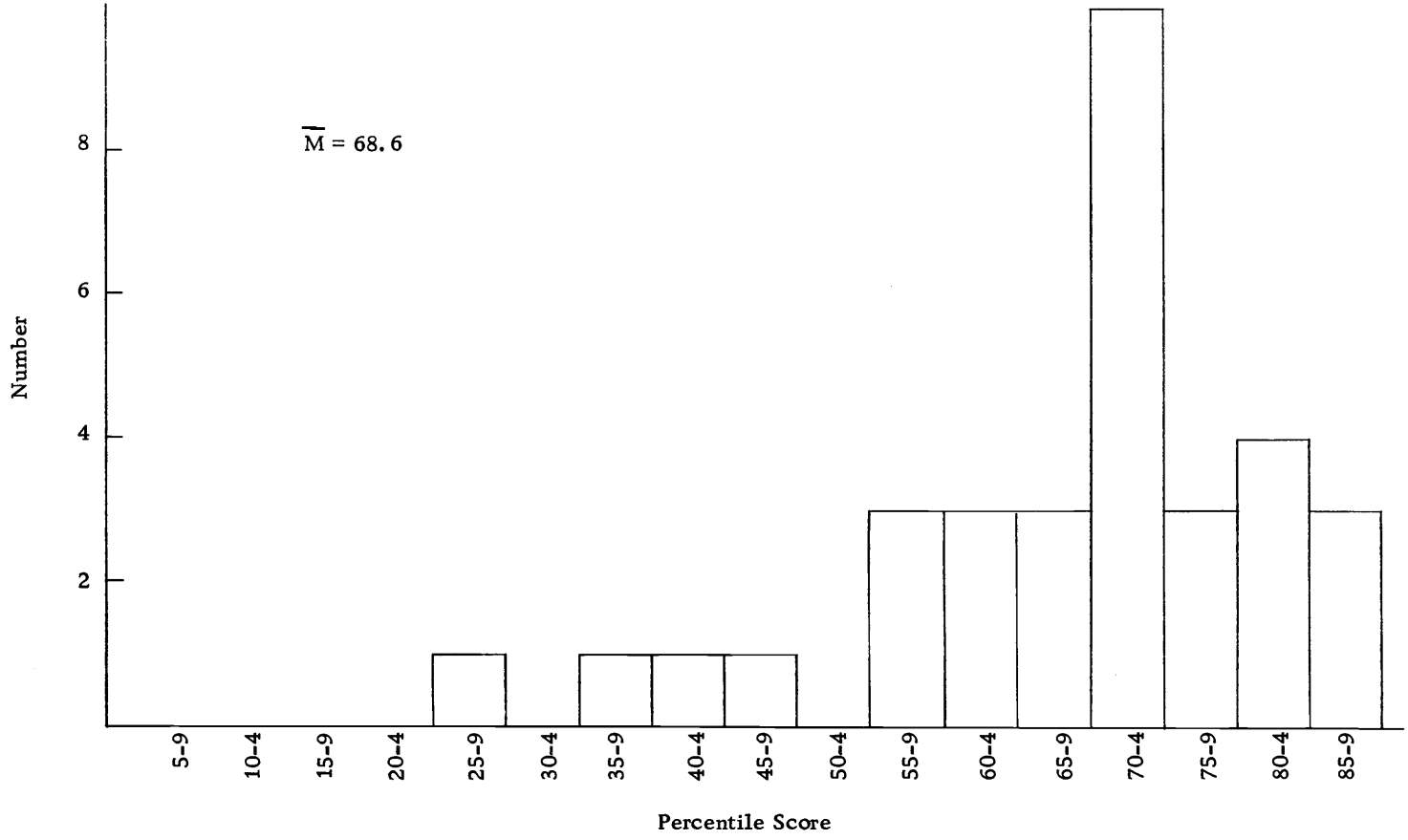

Figure 26. Distribution of Institutional Means on Pilot Aptitude Composite--Final Administration.

found to be .78, somewhat less than was found using Officer Quality Composite ranks at the two administrations. It may well be that this correlation coefficient would be substantially higher were it not for the variation among institutions in their proportion of Category IP cadets. In any case there appears a substantial tendency for cadets of an institution to retain a certain rank-order position among institutions as regards their average Pilot Aptitude Composite scores during their collegiate careers.

As has been previously indicated, a considerable amount of unpublished research has clearly established that the mean Officer Quality Composite score at an institution tends to remain relatively stable over a period of years and that these characteristic means vary over a wide range. Similar research has not been performed using Pilot Aptitude Composite scores. Such research seems warranted.

Figures 27 and 28 portray respectively the distribution of institutional means on the Navigator/Technical Composite for the two administrations. It is immediately apparent that the average improvement in percentile score from the initial to the final administration is very small (5.6) when compared with institutional improvements registered in the other two composites (16.7 and 23.8). This, of course, follows as a matter of course from the low average improvement in individual scores (5.7) for the total group.

The correlation coefficient between institutions' rank orders in

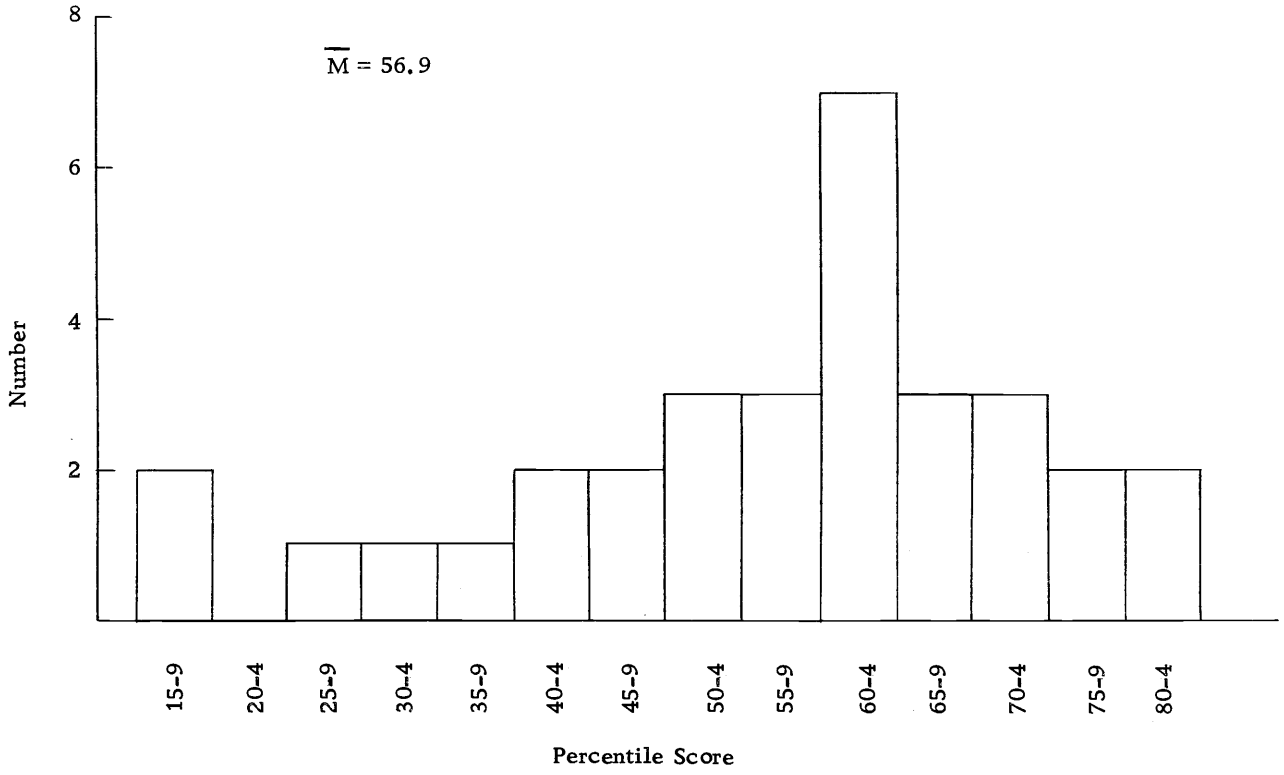

Figure 27. Distribution of Institutional Means on Navigator/Technical Composite--Initial Administration.

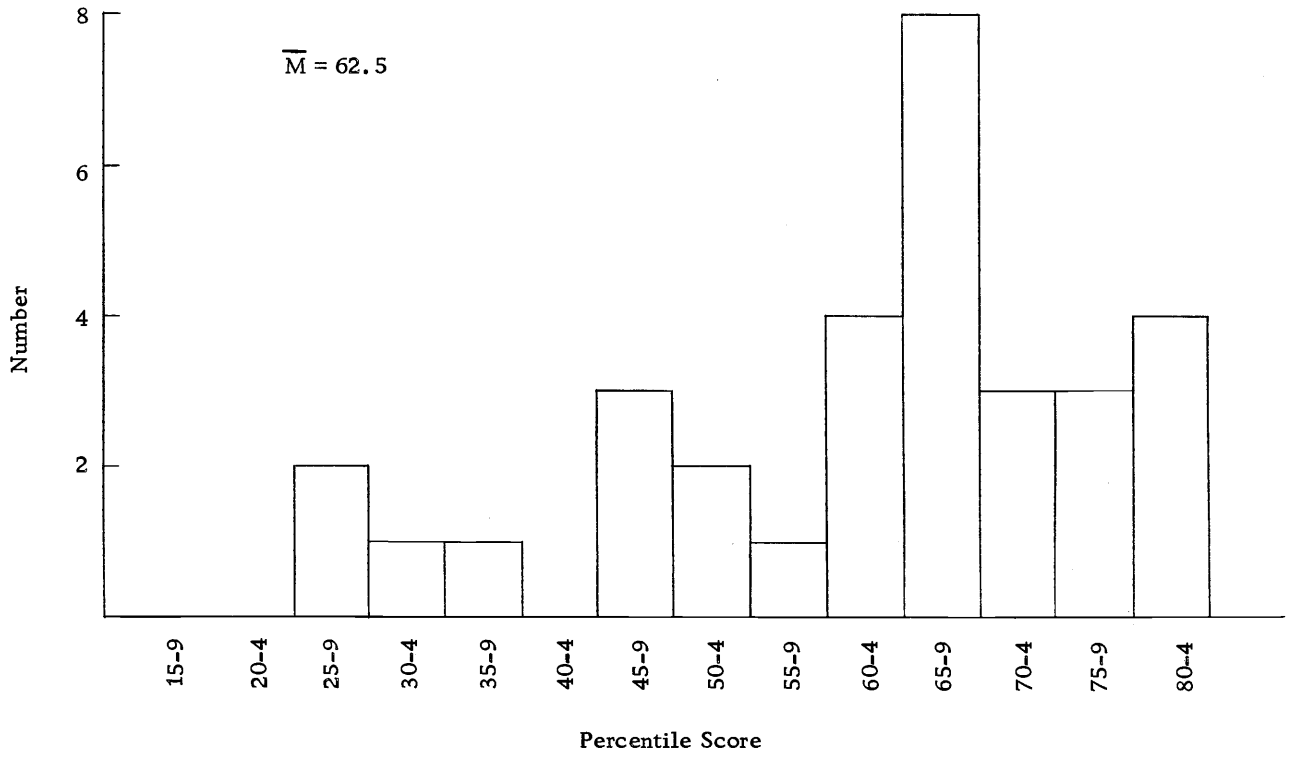

Figure 28. Distribution of Institutional Means on Navigator/Technical Composite--Final Administration.

the Navigator/Technical Composite at the two administrations was .91. It is thus apparent that only very slight changes in rank order have taken place between the two administrations.

Such a high degree of correlation is not surprising. Even a cursory examination of the data from both administrations reveals that, in general, predominately technical and engineering schools rank high on the Navigator/Technical Composite, primarily liberal arts schools are low and universities representing a composite of many disciplines are near the middle, in both administrations.

To ascertain the degree to which institutional ranks on the composites tend to agree, coefficients of correlation were computed among the three composites using rank order of institutional means at the initial administration. These coefficients are presented in Table 5.

Table 5. Correlation Coefficients of Institutional Ranks in the Composites at the Initial Administration

	Composite	
Composite	Officer Quality	Pilot Aptitude
Pilot Aptitude	0.44	
Navigator/Technical	0.72	0.74

These results indicate that there is a substantial tendency for the cadets of an institution to be at about the same relative position

among cadets from other institutions in the Officer Quality and Navigator/Technical Composites and in the Pilot Aptitude and Navigator/Technical Composites. There is little relationship between their relative position on the Officer Quality and Pilot Aptitude Composites.

Reliability of Composites

The fact that individual scores are available for a representative sample of over 400 AFROTC cadets on all three composites from two widely spaced administrations, encourages the computation of correlation coefficients between scores obtained at the two administrations for each composite separately. The results of these computations are shown in Table 6. Scatter diagrams are seen in Tables 7 through 9.

Table 6. Correlation Coefficients Between Individuals' Test Scores at the Two Administrations, for Each Composite

Officer Quality Composite	0.84
Pilot Aptitude Composite	0.71
Navigator/Technical Composite	0.90

In a sense these coefficients are a measure of the reliability of the composites for they indicate the degree to which cadets tended to retain at the final administration, and on a different test form, the same relative position they held at the initial administration.

Table 7. Correlation of Officer Quality Composite Percentile Scores - Initial and Final Administration

Percentile Score - Initial	Percentile Score - Final Administration										
	0-9	10-9	20-9	30-9	40-9	50-9	60-9	70-9	80-9	90-9	Sum
90-9								2	2	36	40
80-9							3	5	16	52	76
70-9						3	6	9	10	25	53
60-9					1	2	6	11	11	24	55
50-9					1	5	6	8	8	8	36
40-9					2	5	6	11	9	7	40
30-9		1			5	4	4	8	5	5	32
20-9			6	5	4	7	7	7			36
10-9	1	6	6	7	6	4	7				37
0-9	2	2	2	1	2	1					10
Sum	3	9	14	13	21	31	45	61	61	157	415

Table 8. Correlation of Pilot Aptitude Composite Percentile Scores - Initial and Final Administration

Percentile Score - Initial	Percentile Score - Final Administration										
	0-9	10-9	20-9	30-9	40-9	50-9	60-9	70-9	80-9	90-9	Sum
90-9										27	27
80-9						1	1		3	19	24
70-9						1	2	2	14	32	51
60-9				1	1	2	7	8	19	18	56
50-9				3	3	1	4	6	10	11	38
40-9				4	7	9	4	5	10	8	47
30-9		1		3	4	3	5	8	5	4	33
20-9		2	1	7	6	6	8	10	7	5	52
10-9		4	8	6	1	4	9	3	3	2	40
0-9	12	12	5	2	6	1	1	4	2	2	47
Sum	12	19	14	26	28	28	41	46	73	128	415

Table 9. Correlation of Navigator/Technical Composite Percentile Scores - Initial and Final Administration

Percentile Score - Initial	Percentile Score - Final Administration										
	0-9	10-9	20-9	30-9	40-9	50-9	60-9	70-9	80-9	90-9	Sum
90-9									19	65	84
80-9						1		7	15	16	39
70-9						1	5	9	24	6	45
60-9						7	8	7	16	2	40
50-9				4	5	9	12	9	7		46
40-9				6	6	11	7	6	1		37
30-9		1	7	8	10	6	4	1			37
20-9	1	5	7	10	6	6	1	1			37
10-9	3	9	7	7	4						30
0-9	8	5	5	2							20
Sum	12	20	26	37	31	41	37	40	82	89	415

Intercorrelation Between Composites

If each of the three composites is truly measuring a unique psychological trait then there should be little or no relationship between scores achieved by individuals on one composite and their scores on either the other two composites. To ascertain the degree to which the composites are in fact associated, correlation coefficients were computed between pairs of composites using test scores obtained at the initial administration and again using test scores obtained at the final administration. The scatter diagrams are presented in Tables 11 through 16. The correlation coefficients are summarized in Table 10.

Table 10. Correlation Coefficients of Individuals' Scores Between Composites

	Initial Administration	
	Officer Quality	Pilot Aptitude
Pilot Aptitude	0.34	
Navigator/Technical	0.76	0.59
	Final Administration	
	Officer Quality	Pilot Aptitude
Pilot Aptitude	0.39	
Navigator/Technical	0.77	0.63

In both administrations the highest degree of correlation has been found to exist between the Officer Quality and Navigator/Technical Composites. This substantial relationship is easily explained by the facts that (1) these composites have more than 25% of their items in common, and (2) both are heavily weighted with items of an academic nature. The low correlation coefficient obtained between Officer Quality and Pilot Aptitude Composite scores demonstrates the relative uniqueness of the traits being measured by the two composites. The moderate relationship found between Pilot Aptitude Composite scores and Navigator/Technical Composite scores may perhaps be explained by a tendency for those with pilot interest to also possess an interest in mechanical and technological matters.

Table 11. Correlation Between Officer Quality and Pilot Aptitude Composites - Initial Administration

Pilot Aptitude	Officer Quality 0-9	10-9	20-9	30-9	40-9	50-9	60-9	70-9	80-9	90-9	Sum
90-9		1			2	2	3	3	8	8	27
80-9		1		2	2	1	3	3	7	5	24
70-9		5	2	2	4	3	6	6	16	7	51
60-9	1	2	4	2	8	7	11	7	10	4	56
50-9	1	3	1	4	5	2	5	7	7	3	38
40-9		2	1	4	4	6	9	6	11	4	47
30-9		1	5	2	8	3	6	4	3	1	33
20-9	1	2	11	4	5	6	4	9	7	3	52
10-9	2	8	4	7	1	4	2	3	6	3	40
0-9	5	12	8	5	1	2	6	5	1	2	47
Sum	10	37	36	32	40	36	55	53	76	40	415

Table 12. Correlation Between Officer Quality and Pilot Aptitude Composites - Final Administration

Pilot Aptitude	Officer Quality 0-9	10-9	20-9	30-9	40-9	50-9	60-9	70-9	80-9	90-9	Sum
90-9	1	0	2	1	2	6	7	17	21	71	128
80-9		1	1	1	3	4	8	10	16	29	73
70-9			1	1		5	7	7	7	18	46
60-9			2	1	2	4	8	9	5	10	41
50-9		2	3	1	2	1	4	4	2	9	28
40-9	1	1	1		4	5	3	3	5	5	28
30-9		1	2	3	4	2	3	3	3	5	26
20-9				2	1	1	2	3	1	4	14
10-9	1	1	1	1	3	3	2	3	1	3	19
0-9		3	1	2			1	2		3	12
Sum	3	9	14	13	21	31	45	61	61	157	415

Table 13. Correlation Between Pilot Aptitude and Navigator/Technical Composites - Initial Administration

Pilot Aptitude	Navigator/Technical 0-9	10-9	20-9	30-9	40-9	50-9	60-9	70-9	80-9	90-9	Sum
90-9				1			2		3	21	27
80-9				1	1	1	3	3	2	13	24
70-9			2	6	2	4	5	7	8	17	51
60-9		1	3	3	5	10	9	8	8	9	56
50-9		3		6	2	3	5	8	5	6	38
40-9		3	4	2	8	7	4	11	3	5	47
30-9	1	2	5	3	6	3	4	5	2	2	33
20-9	1	7	5	3	7	11	6	2	7	3	52
10-9	5	4	9	3	3	7	6	2	1		40
0-9	13	9	8	8	4	1	3	1			47
Sum	20	29	36	36	38	47	47	47	39	76	415

Table 14. Correlation Between Pilot Aptitude and Navigator/Technical Composites - Final Administration

Pilot Aptitude	Navigator/Technical 0-9	10-9	20-9	30-9	40-9	50-9	60-9	70-9	80-9	90-9	Sum
90-9			2	2	2	8	10	10	32	62	128
80-9		1	1	6	6	12	9	5	15	18	73
70-9	1		1	4	3	4	6	11	14	2	46
60-9		3	4	5	7	1	3	4	11	3	41
50-9		3	6	3	1	3	3	2	5	2	28
40-9		3	6	3		5	1	5	4	1	28
30-9	1	2	3	5	4	4	3	2	1	1	26
20-9		2	1	6		2	2	1			14
10-9	4	4	1	2	6	2					19
0-9	6	2	1	1	2						12
Sum	12	20	26	37	31	41	37	40	82	89	415

Table 15. Correlation Between Officer Quality and Navigator/Technical Composites - Initial Administration

Officer Quality	Navigator/Technical 0-9	10-9	20-9	30-9	40-9	50-9	60-9	70-9	80-9	90-9	Sum
90-9						2	3	5	9	21	40
80-9		1	2	6	7	6	7	15	7	25	76
70-9		1		6	6	8	10	8	6	8	53
60-9	1		7	4	6	6	5	8	7	11	55
50-9		2		8	4	5	6	3	2	6	36
40-9	1	2	6	1	7	6	7	4	4	2	40
30-9	3	3	5	4	3	4	3	3	2	2	32
20-9	3	8	6	3	3	8	2	1	1	1	36
10-9	9	11	5	4	2	2	3		1		37
0-9	3	1	4	1			1				10
Sum	20	29	35	37	38	47	47	47	39	76	415

Table 16. Correlation Between Officer Quality and Navigator/Technical Composites - Final Administration

Officer Quality	Navigator/Technical 0-9	10-9	20-9	30-9	40-9	50-9	60-9	70-9	80-9	90-9	Sum
90-9				1	5	10	8	19	47	67	157
80-9			1	3	6	5	8	8	17	13	61
70-9		1	4	7	3	7	7	9	15	8	61
60-9		3	3	8	6	10	10	2	2	1	45
50-9		4	2	7	8	3	4	2	1		31
40-9	1	4	6	3	3	4					21
30-9		2	5	5		1					13
20-9	5	2	3	3		1					14
10-9	3	4	2								9
0-9	3										3
Sum	12	20	26	37	31	41	37	40	82	89	415

A coefficient of determination (the square of the correlation coefficient) may be converted into a percentage statistic. Its application here yields the data in Table 17.

Table 17. Percent of Determination Between Composites

	Officer Quality	Pilot Aptitude
Pilot Aptitude	10-15%	
Navigator/Technical	55-60%	35-40%

These data may be interpreted as follows: e. g., 10 to 15 percent of the variance in scores obtained on the Pilot Aptitude Composite is associated with the variance in the psychological trait measured by the Officer Quality Composite. More than half the variance in Navigator/Technical Composite scores is associated with variance in Officer Quality Composite scores.

Statistical Test of Significance

Chi square for 2 by k tables were computed for figures four through fifteen and figures seventeen through twenty-eight (see Appendix). The formula:

$$X^2 = \frac{N^2}{A_t B_t} \left[\Sigma \frac{B^2_i}{Ai + Bi} - \frac{B^2_t}{A_t + B_t} \right]$$ was used (14).

These computations were made at the Statistics Department of Oregon State University by Mr. James Daley of that department.

This method of computation for significance was approved by William R. Crooks, Chairman, Psychology Department, Oregon State University.

CHAPTER V

CONCLUSIONS

Based solely on the data analyzed in the research reported herein, the following conclusions seem warranted:

1. Three to four years of collegiate experience coupled with the concomitant maturation have a marked elevating effect on scores obtained in the Air Force Officer Qualifying Test.

2. On the average, college seniors can be expected to score a minimum of 30 percentile points higher than college freshmen on the Officer Quality Composite.

3. Category III cadets do not elevate their Officer Quality Composite scores, during collegiate experience, as much as do cadets in the other categories.

4. Percentile scores, on the Pilot Aptitude Composite, of cadets with no known interest or experience in flying (Categories II and III) increase by as much as 20 points, on the average, between freshman and senior year. There seems to be no difference, in this regard, between Category II and III cadets.

5. Intervening flying training between administrations of the AFOQT results in average increases in Pilot Aptitude Composite percentile scores of 30 to 50 points.

6. Category II cadets score significantly higher than Category IP or IN cadets (10 to 15 percentile points higher) on the Navigator/Technical Composite of the AFOQT both as freshmen and as seniors. Category II cadets average a minimum of 30 percentile points higher than Category III cadets on this same composite, both as freshmen and as seniors.

7. Category II cadets who as freshmen score below the 75th percentile on the Navigator/Technical Composite, are likely to average scores about 30 percentile points higher after three to four years of collegiate experience.

8. Not only do successive classes of an institution tend to have a relatively stable, characteristic mean score on the Officer Quality Composite (as has been demonstrated elsewhere) but it appears from this study that each group tends to retain its relative position among institutions throughout the collegiate careers of its members.

9. The Officer Quality and Navigator/Technical Composites of the AFOQT have a high degree of reliability. Cadets tend to retain their relative standing among their peers.

10. The Officer Quality Composite and the Pilot Aptitude Composite appear to be measuring relatively unique psychological traits. There is not more than about 15 percent of determination between them.

11. There is substantial relationship between scores obtained

by individuals on the Officer Quality Composite and the Navigator/Technical Composite (55-60% determination). This relationship stems from the facts that (1) the two composites have 27% of their items in common and (2) both are primarily academic in nature.

12. There is moderate relationship (35-40% determination) between scores achieved by individuals on the Pilot Aptitude Composite and the Navigator/Technical Composite of the Air Force Officer Qualifying Test.

BIBLIOGRAPHY

1. Anastasi, Anne and John P. Foley, Jr., Differential psychology. Rev. ed. New York, Macmillan, 1949. 894 p.

2. Barzun, Jacques and Henry F. Graff. The modern researcher. New York, Harcourt, Brace and World, 1959. 367 p.

3. Best, John W. Research in education. Englewood Cliffs, New Jersey, Prentice-Hall, 1959. 306 p.

4. Carroll, John B. Words, meanings and concepts. Part I. Their nature. Harvard Educational Review 34:178-190. 1964.

5. Christal, Raymond E. and John D. Krumboltz. Use of The Air Force Officer Qualifying Test in The AFROTC Selection Program. Air Force Personnel and Training Research Center, Lackland Air Force Base, Texas. Feb., 1957. A report. 35 p.

6. Coleman, James C. Personality dynamics and effective behavior. Chicago, Scott, Foresman, 1964. 548 p.

7. Cronbach, Lee J. Essentials of psychological testing. 2d ed. New York, Harper and Row, 1960. 650 p.

8. Droege, Robert C. GATB longitudinal maturation study. Personnel and Guidance Journal 44:919-930. 1966. (Reprint)

9. Ebel, Robert L. Measuring educational achievement. Englewood Cliffs, New Jersey, Prentice Hall, 1965. 469 p.

10. English, Horace B. and Ava C. A comprehensive dictionary of psychological and psychoanalytic terms. 3d ed. New York, Green, 1959. 594 p.

11. Garrett, Henry E. Testing for teachers. New York, American Book, 1959. 258 p.

12. Hilgard, Ernest R. Introduction to psychology. 3d ed. New York, Harcourt, Brace and World, 1962. 678 p.

13. Horrocks, John E. The psychology of adolescence behavior and development. 2d ed. Boston, Houghton, Mifflen, 1959. 690 p.

14. Mendenhall, William. Introduction to probability and statistics. Belmont, California, Wadsworth, 1967. 393 p.

15. Nunnally, Jum C., Jr. Tests and measurements. New York, McGraw Hill, 1959. 446 p.

16. Terman, Lewis M. and Maude A. Merrill. Measuring intelligence. Boston, Houghton, 1937. 461 p.

17. Thorndike, Robert L. and Elizabeth Hagen. Measurement and evaluation in psychology and education. 2d ed. New York, Wiley, 1961. 602 p.

18. Tugman, Robert C., Academic achievement and army qualification test scores compared on a basis of 8000 cases. Master's thesis. Corvallis, Oregon State University, 1955. 125 numb. leaves.

19. U. S. Air Force, Air Force personnel evaluation manual. Washington, D. C., 1968. 104 p. (Air Force Manual AFM 35-8). (19, p. 63-66)

20. Weschsler, David. The measurement of adult intelligence. 3d ed. Baltimore, Williams, 1944. 258 p.

APPENDIX

APPENDIX

Frequency Comparison: Chi Square for Figures 4 and 5.

	A	B	C	B/C	B^2/C
	10	3	13	.2308	.692
	37	9	46	.1957	1.761
	36	14	50	.2800	3.920
	32	13	45	.2889	3.756
	40	21	61	.3443	7.230
	36	31	67	.4627	14.343
	55	45	100	.4500	20.250
	53	61	114	.5351	32.640
	76	61	137	.4453	27.161
	40	157	197	.7970	125.122
(Σ)	415	415	830	4.0298	236.875
					207.500
					29.375

$$\chi^2 = \frac{688900}{172225} \qquad 4.00 \times 29.375$$

$n = 117.5$ $p < .001$ Very significant

Frequency Comparison: Chi Square for Figures 6 and 7.

	A	B	A + B	B/A + B	$B^2/A + B$
	4	0	4	0	0
	6	0	6	0	0
	14	4	18	.222	.444
	20	6	26	.231	1.384
	45	7	52	.135	.942
	48	9	57	.158	1.421
	62	13	75	.200	2.253
	47	13	60	.217	2.817
	26	8	34	.235	1.882
	26	14	40	.350	4.900
	33	19	52	.365	7.326
	28	20	48	.417	8.333
	14	8	22	.363	2.909
	13	13	26	.500	6.500
	14	14	28	.500	7.000
	2	4	6	.667	2.667
	7	7	14	.500	3.500
	2	2	4	.500	1.000
	1	1	2	.500	0.500
	2	2	4	.500	1.000
(Σ)	414	164	578		56.778
$(\Sigma)^2$		26896	334084		46.532
					10.246

$$\chi^2 \quad \frac{334084}{414(164)} \quad (10.246) = 50.415$$

$n = 19$ $\quad p < .001$

very significant

Frequency Comparison: Chi Square for Figures 8 and 9.

	A	B	A + B	B/A + B	$B^2/A + B$
	28	5	23	5/23	1.087
	18	6	24	6/24	1.500
	20	10	30	10/30	3.333
	20	8	28	8/28	2.286
	24	4	28	4/28	0.571
	28	8	36	8/36	1.777
	15	18	33	18/33	11.030
	18	8	26	8/26	2.462
	24	16	40	16/40	6.400
	22	12	34	12/34	4.235
	20	12	32	12/32	4.500
	18	15	33	15/33	6.818
	28	18	46	18/46	7.913
	28	21	49	21/49	9.000
	21	22	43	22/43	11.256
	30	22	52	22/52	9.308
	12	35	47	35/47	26.063
	11	38	49	38/49	29.469
	16	48	64	48/64	36.000
	11	80	91	80/91	70.329
(Σ)	412	406	818		245.337
$(\Sigma)^2$		164836	669124		201.510
					43.827

$$\chi^2 = \frac{669124}{412(406)} (43.827) = 175.308$$

$p < .001$

very significant

Frequency Comparison: Chi Square for Figures 10 and 11.

	A	B	A + B	B/A + B	$B^2/A + B$
	24	2	26	2/26	.154
	15	4	19	4/19	.842
	14	7	21	7/21	2.333
	10	7	17	7/17	2.882
	10	8	18	8/18	3.555
	10	13	23	13/23	7.348
	8	7	15	7/15	3.226
	8	10	18	10/18	5.555
	6	13	19	13/19	8.895
	12	10	22	10/22	4.545
	9	10	19	10/19	5.263
	7	8	15	8/15	4.266
	13	15	28	15/28	8.036
	9	14	23	14/23	8.522
	5	10	15	10/15	6.666
	16	14	30	14/30	6.533
	5	8	13	8/13	4.923
	4	7	11	7/11	4.455
	7	7	14	7/14	3.500
	5	4	9	4/ 9	1.777
(Σ)	197	178	375		93.316
$(\Sigma)^2$		31684	140625		84.490
					8.826

$$\chi^2 = \frac{140625}{197(178)} (8.826) = 35.392$$

n = 19 p = .02

very significant

Frequency Comparison: Chi Square for Figures 12 and 13.

	A	B	A + B	B/A + B	$B^2/A + B$
	5	0	5	0	0
	7	0	7	0	0
	9	0	9	0	0
	6	0	6	0	0
	4	0	4	0	0
	6	1	7	1/ 7	.143
	12	2	14	2/14	.286
	6	3	9	3/ 9	1.000
	11	3	14	3/14	.643
	9	1	10	1/10	.100
	8	4	12	4/12	1.333
	8	7	15	7/15	3.266
	11	7	18	7/18	2.722
	8	13	21	13/21	8.048
	10	10	20	10/20	5.000
	14	10	24	10/24	4.166
	13	15	28	15/28	8.036
	14	20	34	20/34	11.765
	14	32	46	32/46	22.260
	22	52	74	52/74	36.540
(Σ)	197	180	377		105.308
$(\Sigma)^2$		32400	142129		85.941
					19.367

$$\chi^2 = \frac{142129}{197(180)} (19.367) = 77.622 \qquad n = 19$$

$p < .001$

very significant

Frequency Comparison: Chi Square for Figures 14 and 15.

A	B	A + B	B/A + B	$B^2/A + B$
1	0	1	0	0
1	2	3	2/3	1.333
0	3	3	3/3	3.000
4	1	5	1/5	.200
3	3	6	3/6	1.500
2	3	5	3/5	1.800
12	16	28	16/28	9.143
11	12	23	12/23	6.261
5	12	17	12/17	8.471
12	10	22	10/22	4.545
8	10	18	10/18	5.555
6	3	9	3/9	1.000
6	5	11	5/11	2.273
4	4	8	4/8	2.000
7	5	12	5/12	2.083
3	3	6	3/6	1.500
4	4	8	4/8	2.000
2	2	4	2/4	1.000
0	1	1	1	1.000
0	1	1	1	1.000
1	0	1	0	0
0	2	2	2/2	2.000
1	1	2	1/2	.500
0	0	0	0	0
0	1	1	1	1.000
(Σ) 93	104	197		59.164
$(\Sigma)^2$	10816	38809		54.903
				4.261

$$\chi^2 = \frac{38809}{(104)93} (4.261) = 17.099$$

n = 24

p = .50 not significant

Frequency Comparison: Chi Square for Figures 17 and 18.

	A	B	A + B	B/A + B	$B^2/A + B$
	8	2	10	2/10	.400
	12	10	22	10/22	4.545
	12	11	23	11/23	5.261
	18	9	27	9/27	3.000
	21	13	34	13/34	4.970
	15	13	38	13/38	4.447
	20	20	40	20/40	10.000
	17	17	34	17/34	8.500
	21	19	40	19/40	9.525
	16	12	28	13/28	5.143
	21	17	38	17/38	7.605
	25	24	49	24/49	11.755
	23	19	42	19/42	9.071
	17	17	34	17/34	8.500
	21	21	42	21/42	10.500
	24	19	43	19/43	8.860
	17	43	60	43/60	30.816
	21	39	60	39/60	25.350
	34	57	91	57/91	35.703
	50	32	82	32/82	12.487
(Σ)	413	424	837		216.434
$(\Sigma)^2$		179776	700569		214.780
					2.658

$$\chi^2 = \frac{700569}{413(424)} (2.658) = 10.623$$

n = 19

p = .95 not significant

Frequency Comparison: Chi Square for Figures 19 and 20.

A	B	A + B	B/A + B	$B^2/A + B$
6	1	7	1/7	.143
8	9	17	9/17	4.765
6	5	11	5/11	2.273
5	4	9	4/9	1.778
7	6	13	6/13	2.769
6	6	12	6/12	3.000
6	7	13	7/13	3.769
6	8	14	8/14	4.571
9	6	15	6/15	2.400
6	4	10	4/10	1.600
4	7	11	7/11	4.455
6	4	10	4/10	1.600
6	7	13	7/13	3.769
1	3	4	3/4	2.250
4	4	8	4/8	2.000
5	3	8	3/8	1.125
1	7	8	7/8	6.125
4	5	9	5/9	2.778
5	8	13	8/13	4.923
4	0	4	0	0
$(\Sigma)_2$ 105	104	209		56.093
$(\Sigma)^2$	10816	43681		51.751
				4.342

$$\chi^2 = \frac{43681}{105(104)} (4.342) = 17.368$$

n = 19

p = .50 not significant

Frequency Comparison: Chi Square for Figures 21 and 22.

	A	B	A + B	B/A + B	$B^2/A + B$
	0	1	1	1	1.000
	2	0	2	0	0
	1	2	3	.667	1.333
	5	0	5	0	0
	2	0	2	0	0
	3	1	4	.250	.250
	2	4	6	.667	2.667
	1	3	4	.775	2.250
	2	1	3	.333	.333
	2	1	3	.333	.333
	2	1	3	.333	.333
	2	4	6	.667	2.667
	4	3	7	.428	1.286
	5	2	7	.286	.571
	3	5	8	.625	1.250
	8	6	14	.428	2.571
	5	12	17	.706	8.470
	7	14	21	.667	9.333
	14	18	32	.562	11.063
	23	15	38	.395	5.921
	93	93	186		51.631
$(\Sigma)_2$		8649	34596		46.500
$(\Sigma)^2$					5.131

$$\chi^2 = \frac{34596}{93(93)} (5.131) = 20.524$$

n = 19

p > .30 not significant

Frequency Comparison: Chi Square for Figures 23 and 24.

	A	B	A + B	B/A + B	$B^2/A + B$
	1	0	1	0	0
	1	0	1	0	0
	0	0	0	0	0
	1	0	1	0	0
	2	2	4	.5	1.000
	6	0	6	0	0
	5	1	6	.167	.167
	6	3	9	.333	1.000
	0	4	4	1.000	4.000
	4	2	6	.333	.667
	5	4	9	.444	1.778
	1	5	6	.833	4.166
	0	6	6	1.000	6.000
	0	5	5	1.000	5.000
(Σ)	31	32	63		23.778
$(\Sigma)^2$		1024	3969		16.253
					7.525

$$\chi^2 = \frac{3969}{31(32)} (7.525) = 30.100$$

n = 19

p = .05 significant

Frequency Comparison: Chi Square for Figures 25 and 26.

	A	B	A + B	B/A + B	B^2/A + B
	1	0	1	0	0
	0	0	0	0	0
	3	1	4	.250	.250
	5	1	6	.167	.167
	11	2	13	.154	.308
	5	3	8	.375	1.125
	3	6	9	.556	4.000
	1	13	14	.929	12.071
	0	7	7	1.000	7.000
(Σ)	29	33	62	3.431	24.921
					17.565
					7.356

$62^2 = \frac{3844}{957}$ 4.017 x 7.356

$\chi^2 = 29.589$

n = 9

p < .001 very significant

Frequency Comparison: Chi Square for Figures 27 and 28.

	A	B	A + B	B/A + B	$B^2/A + B$
	2	0	2	0	0
	1	2	3	.667	1.333
	2	2	4	.500	1.000
	4	3	7	.429	1.286
	6	3	9	.333	1.000
	10	12	22	.545	6.545
	5	6	11	.545	3.273
	2	4	6	.667	2.667
(Σ)	32	32	64	3.686	17.104
					16.000
					1.104

$\frac{64^2}{32 \times 32} = \frac{4096}{1024} = 4$ $\quad 4 \times 1.104 \quad \chi^2 = 4.416$

n = 9

p = .50 not significant