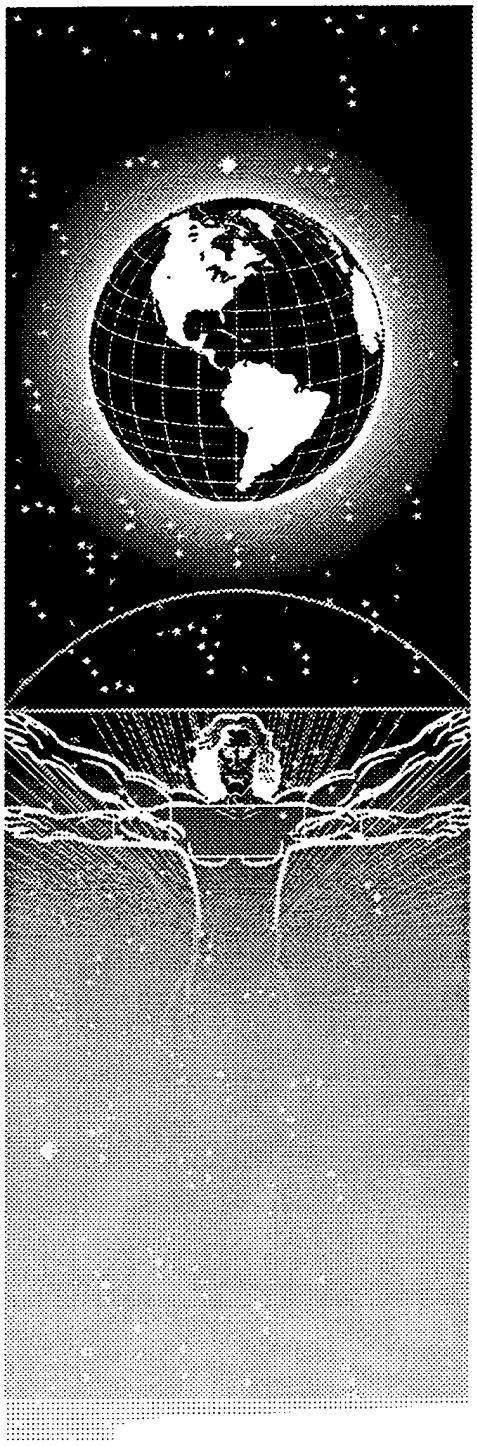
AL/HR-TP-1996-0036



UNITED STATES AIR FORCE ARMSTRONG LABORATORY

AIR FORCE OFFICER QUALIFYING TEST (AFOQT): FORMS Q DEVELOPMENT, PRELIMINARY EQUATING AND OPERATIONAL EQUATING

Theresa M. Glomb

University of Illinois at Urbana-Champaign
Department of Psychology
603 East Daniel
Champaign, IL 61820

James A. Earles

HUMAN RESOURCES DIRECTORATE
MANPOWER AND PERSONNEL RESEARCH DIVISION
7909 Lindbergh Drive
Brooks Air Force Base, TX 78235-5352

July 1997

AIR FORCE MATERIEL COMMAND
ARMSTRONG LABORATORY
HUMAN RESOURCES DIRECTORATE
7909 Lindbergh Drive
Brooks Air Force Base, TX 78235-5352

NOTICES

Publication of this paper does not constitute approval or disapproval of the ideas or findings. It is published in the interest of scientific and technical information (STINFO) exchange.

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility or any obligation whatsoever. The fact that the Government may have formulated or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication, or otherwise in any manner construed, as licensing the holder, or any other person or corporation; or as conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The Office of Public Affairs has reviewed this paper, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This paper has been reviewed and is approved for publication.

JAMES A. EARLES
Project Scientist

PATRICK C. KYLLONEN
Technical Director

GARY D. ZANK, Colonel, USAF
Chief, Manpower & Personnel Research Division

Please notify this office, AL/HRPP, 7909 Lindbergh Drive, Brooks AFB TX 78235-5352, if your address changes, or if you no longer want to receive our technical reports. You may write or call the STINFO office at DSN 240-3853 or commercial (210) 536-3853

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 1997	Final - January 1991-August 1996

4. TITLE AND SUBTITLE
Air Force Officer Qualifying Test (AFOQT): Forms Q Development, Preliminary Equating and Operational Equating

5. FUNDING NUMBERS
PE - 62205F
PR - 1123
TA - A3
WU - 05

6. AUTHOR(S)
Theresa M Glomb
James A Earles

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Armstrong Laboratory
Human Resources Directorate
Manpower & Personnel Research Division
7909 Lindbergh Drive
Brooks AFB TX 78235-5352

8. PERFORMING ORGANIZATION REPORT NUMBER
AL/HR-TP-1996-0036

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
Armstrong Laboratory Technical Monitor: James A. Earles, (210) 536-3257

12a. DISTRIBUTION AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
Air Force Officer Qualifying Test (AFOQT) Forms Q1 and Q2 were developed to be highly similar to the just previous Forms P1 and P2 in test content and length, item difficulty and discrimination, and stylistic features. Approximately half the items were new and half were from Forms P. Forms Q1 and Q2 were preliminarily equated on Air Force Academy students, Reserve Officer Training Corps cadets and Officer Training School cadets. Scoring tables were developed and AFOQT Forms P1, Q1, and Q2 were administered operationally to over 3000 cases per form. Equating tables were accomplished using the data from these actual applicants. It was concluded that these scoring tables should be implemented to make Air Force Officer selection decisions.

14. SUBJECT TERMS
Air Force Officer Qualify Test
AFOQT
Equating

15. NUMBER OF PAGES
80

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

TABLE OF CONTENTS

TABLE OF CONTENTS (CONCLUDED)

LIST OF TABLES

LIST OF TABLES

AIR FORCE OFFICER QUALIFYING TEST (AFOQT): FORMS Q DEVELOPMENT, PRELIMINARY EQUATING AND OPERATIONAL EQUATING

INTRODUCTION

The Air Force Officer Qualifying Test (AFOQT) provides aptitude measures for the Air Force's officer selection system. The AFOQT is used to select individuals for Officer Training School, to select Reserve Officer Training Corps (ROTC) cadets for the Professional Officers Training Course and scholarships, and to select students for Undergraduate Pilot Training and Undergraduate Navigator Training. Air Force Academy applicants are not required to take the AFOQT prior to entry.

The forerunners of the AFOQT were the Aviation Cadet Qualifying Examination (ACQE), a general abilities screening test that was later replaced by the Aviation Cadet Qualifying Test (ACQT) , and the Aircrew Classification Battery (ACB), used to screen and classify potential aircrew members. A preliminary version of the AFOQT was developed by 1952 and the AFOQT Form A replaced the previous test batteries by 1955. A more comprehensive account of the history and development of the AFOQT testing program was authored by Rogers, Roach and Short, 1986. Since the initial Form A, the AFOQT has been updated periodically to ensure currency, security and predictive validity. Forms Q are the seventeenth versions of the AFOQT.

The periodic updates of the AFOQT have historically been the responsibility of the Air Force Human Resources Laboratory (AFHRL) now the Human Resources Directorate of the Air Force's Armstrong Laboratory. Updating the AFOQT currently begins

with a test development of parallel test forms that are equivalent to previous AFOQT test forms on item specifications such as statistics and content. In addition to the test development process, updating the AFOQT involves a provisional equating and operational equating.

The purpose of this report is to describe the construction of the AFOQT Forms Q1 and Q2 and the subsequent equating of these forms to the previous Forms P. The first section discusses item selection and the procedures involved in constructing Forms Q. The second section covers the item, subtest and composite level statistics, and equating statistics of the 1993 data collection used for the preliminary equating analyses. The third section provides this information for the 1995 data used in the operational equating analyses. These equating analyses are integral in linking the new forms of the AFOQT to previous forms to ensure equivalence of measurement.

DEVELOPMENT OF FORMS Q

This test development project began in January of 1991 with the objective of developing two new AFOQT forms, Q1 and Q2, that would be equivalent with previous Forms O, P1 and P2. In maintaining continuity in the testing program, Forms Q were developed to be as similar as possible to previous forms in terms overall test content, test length, item difficulty, item discrimination, subject matter, and stylistic features. Based on prior analyses, the item difficulties of the Form P1 subtests were considered to be more similar to Forms O than P2 item difficulties and P1 was thus chosen as the target test to be replicated.

Test Content

The test content, length, subtests, composite composition and testing time of Forms P remained the same in Forms Q. The AFOQT has 380 items comprising 16 subtests which are combined to create five composite scores. The subtest names, the number of items in each subtest and their categorization into the five composites are presented in Table 1. Total testing time, including administrative procedures, is approximately 270 minutes. A more detailed description of the subtest content can be found in the AFOQT Forms P Test Manual (Berger, Gupta, Berger, & Skinner, 1990).

Table 1. Description of AFOQT Forms Q Subtests and Composition of Aptitude Composites

Subtest		Number of items	Testing time (minutes)	Composites				
				Pilot	Nav-Tech	Acad. Apt.	Verbal	Quant.
Verbal Analogies	(VA)	25	8	X		X	X	
Arithmetic Reasoning	(AR)	25	29		X	X		X
Reading Comprehension	(RC)	25	18			X	X	
Data Interpretation	(DI)	25	24		X	X		X
Word Knowledge	(WK)	25	5			X	X	
Math Knowledge	(MK)	25	22		X	X		X
Mechanical Comprehension	(MC)	20	22	X	X			
Electrical Maze	(EM)	20	10	X	X			
Scale Reading	(SR)	40	15	X	X			
Instrument Comprehension	(IC)	20	6	X				
Block Counting	(BC)	20	3	X	X			
Table Reading	(TR)	40	7	X	X			
Aviation Information	(AI)	20	8	X				
Rotated Blocks	(RB)	15	13		X			
General Science	(GS)	20	10		X			
Hidden Figures	(HF)	15	8		X			
Total		380	208[a]					

Note. [a] This testing time is for minutes actually spent on the test items. Total test time including administrative activities is 270 minutes.

One particularly noteworthy feature of later AFOQT forms is their continuity, which is maintained by the inclusion of anchor or common items in the test forms. The more recent

AFOQT tests have a certain proportion of their items, usually near one-half, that are taken from the previous version, some of which were taken from the version previous to that, and so on. Thus, a subset of the items remains consistent until they are cycled out.

Item Selection

In selecting items for use, comparability was sought between Form P1 and Forms Q in terms of the distributions and mean estimated (or expected) item difficulty for each subtest, the distribution of item content and style, and difficulty associated with the item's position in the subtest. In addition, all item biserial correlations were desired to be positive and high (above .40) for the correct alternative and negative for incorrect alternatives. Both common items and new items were selected to maximize these desired characteristics.

New and Common Items

New items were selected for Forms Q from the same experimental item bank that was used to construct Forms P (Berger, Gupta, Berger, & Skinner, 1988). AFHRL had contracted with Psychometrics, Inc., to create this item bank from which items were to be selected for two new parallel tests, Forms P. The Forms P test development effort marked the first time two equivalent forms of the AFOQT were to be created. After the Forms P test development effort, a sufficient number of items remained in the pool to create Forms Q. Now however, the item bank is depleted and cannot be used for subsequent test construction.

In addition to the new items, approximately half of the items on Forms Q subtests came from Forms P. About half the items in Forms P were also in Form O. Thus, approximately one-quarter of the items are common to all three forms. The same set of common items are in both Forms Q1 and Q2.

Item difficulty and discrimination

The goal of item selection in constructing Forms Q was to match as closely as possible the item difficulties of Form P1, while maintaining the ability to discriminate well between differing levels of ability. The classical item analyses statistics of item difficulty and item discrimination were used to make item selection decisions. For selection of the common items, item statistics from operational use of Forms O and P were used. For selection of the new items, item statistics were obtained from an administration of experimental items from the test bank to a sample that included airmen basic trainees, Reserve Officer Training Corps (ROTC) cadets, and cadets attending Officer Training School (OTS).

The information about items contained in the test bank was based on the experimental test administration to airmen and cadets. The problems of comparing these item statistics from airmen and cadets with those of officer candidates was addressed during the Forms P test development. Multiple regression analyses were conducted to derive weights that could be used to estimate difficulties that would be obtained if items were administered to actual officer candidates. These procedures are discussed in Steuck, Watson, and Skinner (1988).

In development of Forms Q, these same item difficulty estimating techniques were also applied with some exceptions. For all but three subtests, comparisons between the estimated difficulty values for new items selected for Forms P and actual item difficulties obtained during operational use of Forms P produced differences of less than .10 and allowed for the use of item difficulties for Forms P without adjustment. Differences larger than .10 were found on the Data Interpretation, Instrument Comprehension, and Block Counting subtests, for which additional computational adjustments were made to arrive at the expected value. These item difficulty

expected values were obtained with a regression formula which replaced P1 item difficulties with P1 item difficulties reduced by the difference between the mean estimated item difficulty and the mean obtained item difficulty. These adjustments ensured that items selected for Forms Q on these three subtests were approximate to those of Forms P in terms of item difficulty.

Indices of item discrimination, as assessed by the biserial correlation between the item and subtest total were computed for each item. Items were selected in an effort to approximate the distribution of item discrimination values of Form P1. Utilizing the same criteria for construction of previous test forms, items were selected when the correct alternative had a high positive biserial correlation (above .40) and all incorrect alternatives had negative biserial correlations.

Subject Matter

Items from eight of the sixteen subtests are able to be classified according to content categories. The distribution of items within these content categories was to remain consistent across test forms. These content categories were initially constructed in an attempt to match the content of Forms P to that of Forms O. These same classification categories were used for the development of Forms Q with one exception; the Arithmetic Reasoning subtest content categories were modified from the original categories to provide a more empirical framework. Most subtests were categorized on the basis of thematic concerns, others were categorized based on structural or graphic considerations. Content classification strategies were used for the Verbal Analogies, Arithmetic Reasoning, Reading Comprehension, Word Knowledge, Math Knowledge, Scale Reading, Aviation Information and General Science subtests. The remaining subtests were not content classified, generally because all items were essentially the same in content.

Stylistic Features

Stylistic features of Forms O, including format, appearance on a page, type size, type face, illustrations and legend characteristics, consistency of spelling and mathematical notations, were closely matched in Forms P. These same stylistic features were to be replicated as closely as possible in Forms Q, with one exception. The Bodoni typeface of previous AFOQT test forms, a rarely used font, was replaced with the Times Roman font in Forms Q for purposes of printed copy clarity and familiarity.

Following item selection, the subtests were submitted to the monitoring agency to be checked for accuracy, spelling errors, typographical errors, inter-item clueing, distribution of common items, overall presentation, etc. Adjustments were made and replacement items selected where appropriate.

Results and Discussion

The overall results indicate that Forms Q1 and Q2 were closely equivalent with Form P1 and with one another with respect to the item selection criteria. Item difficulty analyses (using estimates of difficulty for officer candidates) suggest that Forms Q1 and Q2 were nearly identical; mean subtest item difficulty values differed by less than .003. In addition, the distributions of item difficulties on both Forms Q1 and Q2 were quite close for most of the subtests, as is evident in Table 2. Comparisons between subtest mean item difficulty of Forms Q and Form P1 further attest to the comparability of forms with respect to item difficulty, as these discrepancies were less than .01 for ten subtests. For the Reading Comprehension, Data Interpretation, Mechanical Comprehension and Block Counting subtests, these discrepancies in subtest mean item difficulty were still small, between .02 and .03.

Item biserial statistics similarly showed a well developed test, with all subtests having mean biserial correlations well over .50. These mean biserial correlations for Forms Q1 and Q2 subtests were similar to one another, as well as to Form P1 subtests, and had no systematic pattern of being higher or lower in one test form. Table 3 shows the similarity of the distributions of biserial correlations for Forms Q1 and Q2. In some selection decisions, item content concerns took precedence over item biserial correlations in order to reduce item redundancy and improve subtest content quality. Some of these lower mean biserial correlations and discrepancies in the distributions for test Forms Q can be traced to substituting items with lower biserial correlations on the basis of content concerns.

Table 2. Distribution of Item Difficulties for Forms Q1 and Q2 at Test Construction Stage

Subtest	Item Difficulties Forms Q1 (Q2)								
	.10-.19	.20-.29	.30-.39	.40-.49	.50-.59	.60-.69	.70-.79	.80-.89	.90-.99
Verbal Analogies	0	1	2	5	3	4	4	3	3
Arithmetic Reasoning	0	0	6	3	2	6	4	4	0
Reading Comprehension	0	0	3	2	4	6	8	2	0
Data Interpretation	0	1	1	3 (2)	6 (9)	7 (5)	7 (6)	0 (1)	0
Word Knowledge	0	0	3	5	5	5	4	3	0
Math Knowledge	0	0	0	2	7	8	6	2	0
Mechanical Comprehension	0	0	4	8	5	3	0	0	0
Electrical Maze	6	3	2	2	3	4	0	0	0
Scale Reading Instrument Comprehension	0	0	4	2	6	6	2	0	0
Block Counting	1 (0)	0	3 (4)	2	3	4 (3)	4 (6)	2 (1)	1
Table Reading									
Aviation Information	0	5	5	5	2	2	1	0	0
Rotated Blocks	0	2	2	5	1	2	0	3	0
General Science	0	2	4	5	4	4	1	0	0
Hidden Figures	0	0	2	3	2	2	2	2	2

Note. [a] Reported values are for Forms Q1 and Q2 except where a parentheses surround the value for Form Q2 indicating a discrepancy.

Table 3. Distribution of Biserial Correlations for Forms Q1 and Q2 at Test Construction Stage

Subtest	Biserial Correlations Forms Q1 (Q2)							
	.20-.29	.30-.39	.40-.49	.50-.59	.60-.69	.70-.79	.80-.89	.90-.99
Verbal Analogies	0 (1)	0	6 (2)	12 (14)	6	1 (2)	0	0
Arithmetic Reasoning	0	0	1 (3)	8 (9)	13 (9)	3 (4)	0	0
Reading Comprehension	0	0	6 (7)	10 (9)	7	2	0	0
Data Interpretation	0	1	8 (6)	9 (12)	6	1 (0)	0	0
Word Knowledge	0	1	3 (1)	6 (9)	11 (10)	4	0	0
Math Knowledge	0	0	3	8 (10)	6 (5)	7 (6)	1	0
Mechanical Comprehension	0	1	4 (2)	6 (5)	8 (9)	1 (3)	0	0
Electrical Maze	0	2	3	5 (4)	8 (5)	2 (6)	0	0
Scale Reading	0	0 (2)	21 (20)	13 (11)	5 (6)	1	0	0
Instrument Comprehension	0	0	0	6 (5)	5	6	3 (4)	0
Block Counting	0	0	4 (2)	3 (5)	5	5 (6)	3 (2)	0
Table Reading	0	3 (0)	4 (7)	7 (5)	9	9 (11)	6 (7)	2 (1)
Aviation Information	0	0	3 (2)	4	10 (8)	2 (4)	1 (2)	0
Rotated Blocks	0	2 (1)	0 (2)	7 (8)	4 (2)	2	0	0
General Science	0	0	7 (5)	9 (10)	4	0 (1)	0	0
Hidden Figures	0	0	0	2	4 (5)	7	2 (1)	0

Note. [a] Reported values are for Forms Q1 and Q2 except where a parentheses surround the value for Form Q2 indicating a discrepancy.

Subject matter comparability between Forms P and Q was achieved to a satisfactory degree. Comparisons between the frequency counts of content categories within a subtest indicate that forms Q1 and Q2 were quite similar; four of the subtests had differing numbers of items per category, none of which were more than two items per category differences. While content category frequency differences were more numerous in comparisons of Forms P with Forms Q than comparing the two Forms Q, the differences overall were still moderate in size.

The stylistic features and format of Forms Q1 and Q2 are equivalent and closely compare to Forms P in most respects, despite the antecedent considerations for item selection. Forms Q correct response options were well balanced across all possible response choices. In some instances, rearrangement of item responses, where rearrangement was not expected to have an impact on examinee performance (some subtests use rules for arranging item options, such as smallest number for option a, next smallest number for option b, etc. and a change would give a clue to the examinee) was necessary to achieve this objective.

In summary, the Forms Q test development effort seems to have achieved the objective of creating two parallel forms that are comparable to Form P1 in terms of item difficulty, item discrimination, subtest content and stylistic considerations. The products of these efforts, the Forms Q1 and Q2, were used in the subsequent test evaluation phase in which preliminary and operational equating tables were developed. These first of these phases, concerned with developing preliminary equatings is discussed next.

PRELIMINARY EQUATING STUDY

Subjects

Subject samples for the preliminary equating study were selected on availability but also to have a broad range of ability. For this purpose, examinees selected were from samples of the Air Force Academy sophomore and junior class, Air Force ROTC cadets, and airmen from the Basic Military Training School. Hereafter these samples will be referred to as AFA, ROTC, and BMTS respectively. ROTC and BMTS examinees were tested from mid-June to mid-August in 1992. The AFA examinees were tested during the end of the school year in 1993.

Demographic information is presented for the total sample and these three subsamples in Table 4. Subjects were predominately male, Caucasian, high school graduates and had attained approximately fourteen or fifteen years of education.

Administration

The AFOQT data for the equating study were collected during four and one-half hour testing sessions during which the standardized test procedures were observed as closely as possible. The standardized procedures for administration are provided in the AFOQT Manual For Administration for Forms Q1 and Q2, a document issued by Air Force Personnel Center (AFPC) that explicates standard test conditions, test material preparation, the use of proctors, and the protocol for conducting the testing session. Testing occurred at Lackland Air Force Base for the examinees from the ROTC and BMTS samples and at the Air Force Academy for AFA examinees.

Table 4. Demographic Percentages for Total, AFA, ROTC and BMTS samples from Preliminary Equating Study

		Total sample			AFA sample			ROTC sample			BMTS sample		
		P1 n=833	Q1 n=810	Q2 n=810	P1 n=217	Q1 n=202	Q2 n=215	P1 n=414	Q1 n=412	Q2 n=394	P1 n=191	Q1 n=188	Q2 n=197
Gender	Male	84.7	84.7	84.8	90.9	92.6	95.4	84.1	83.3	81.4	81.9	80.9	80.8
	Female	15.0	15.2	14.6	8.2	7.4	4.2	15.7	16.5	17.9	18.1	19.1	18.7
	Missing	0.4	0.1	0.6	0.9	0.0	0.5	0.2	0.2	0.8	0.0	0.0	0.5
Race	American Indian	1.3	0.4	0.4	1.4	0.5	0.0	1.2	0.0	0.3	1.6	1.1	1.0
	Asian	2.5	4.6	3.3	1.8	4.5	2.3	3.6	6.1	4.5	1.0	1.6	1.5
	Black	7.3	6.4	8.8	4.6	3.5	3.2	5.8	5.1	9.6	13.0	11.7	13.1
	Hispanic	6.2	4.8	5.8	6.4	3.5	6.0	6.5	4.6	4.8	5.7	6.9	7.6
	Caucasian	82.1	83.6	81.3	85.4	87.6	88.0	82.4	84.0	80.4	77.7	78.7	76.8
	Missing	0.6	0.2	0.4	0.5	0.5	0.5	0.5	0.2	0.5	1.0	0.0	0.0
Years of education	12	16.0	15.5	16.6	0.0	0.0	0.0	0.0	0.0	0.0	66.3	63.8	67.7
	13	3.1	4.2	4.7	0.5	0.0	0.0	0.2	0.7	0.5	11.9	16.0	17.7
	14	49.9	47.0	47.7	39.7	37.1	43.5	72.3	68.3	69.8	15.0	12.8	8.6
	15	23.1	26.3	25.3	49.8	52.5	48.6	19.3	24.7	24.2	2.1	2.7	2.0
	16	5.7	5.7	4.3	9.1	7.9	7.4	5.5	5.3	4.0	2.1	3.7	1.5
	17	1.2	1.0	0.9	0.5	2.5	0.5	1.7	0.5	0.5	1.0	0.5	2.0
	18	0.4	0.2	0.4	0.0	0.0	0.0	0.5	0.5	0.8	0.5	0.0	0.0
	19	0.1	0.1	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.5	0.0
	20	0.1	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	21	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	Missing	0.4	0.0	0.2	0.0	0.0	0.0	0.2	0.0	0.3	1.0	0.0	0.5
Highest degree earned	High School	92.6	93.2	93.1	98.2	100.0	98.1	91.3	90.3	90.7	89.1	92.6	92.4
	Associates	3.8	4.2	4.3	1.4	0.0	1.9	4.8	6.8	6.8	4.7	3.2	2.0
	Bachelors	2.3	1.8	1.8	0.0	0.0	0.0	3.4	1.9	2.0	2.1	3.2	3.5
	Masters	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	Doctorate	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0
	Missing	1.3	0.6	0.7	0.5	0.0	0.0	0.5	0.7	0.6	4.1	1.1	2.0

Data Analysis

The data analysis procedures for both the 1993 Preliminary Equating Study and the 1995 Operational Equating Study were nearly identical. Therefore, the data analysis section will be presented only once for the 1993 Preliminary Equating Study, but will serve for the 1995 Operational Equating Study as well. Variations on this data analysis procedure will be noted where appropriate, however, the major difference is that analyses will be presented for the subgroups of AFA, ROTC, and BMTS (in Appendix A) so that future equating efforts will have the opportunity to inform its data collection from previous efforts.

The subtests and composite formation used in these analyses are as defined previously in Table 1, with two exceptions. After Forms P operational test booklets were printed, two items on test form P1 were determined to be problematic and were not used in subtest scoring; one item from Aviation Information and one item from General Science were omitted. Therefore, the number of items for these two subtests for Form P1 differs by one from those for the corresponding subtests of Forms Q1 and Q2.

Based on item omitting rates and omit patterns, it was determined that two subtests, Scale Reading and Table Reading, should be analyzed as speeded subtests. For these two subtests, the speeded computational formulas for item statistics were used. The remaining subtests were analyzed as power subtests, even though many have a slight speeded component and would probably be correctly classified as mixed-model subtests.

Classical Item Analysis

Item level data were computed using true score theory (Gulliksen, 1950) item statistics such as item difficulties and item discrimination. Item difficulties (p) are defined as the proportion

of examinees who respond correctly to an item. Item difficulties can range from 0.0 to 1.00. Items with difficulties between 0.0 and .30 have a low proportion of respondents answering correctly and are considered hard items. Items with difficulties between .70 and 1.00 have a high proportion of respondents answering correctly and are considered easy items. The reader should note that the term item difficulty is a technical term and seems contradictory to the lay person's definition of difficulty. An item with a low item difficulty is not an item of low difficulty, but rather a very difficult item.

Biserial correlations (rbis), the correlation between the dichotomously scored item and the continuously distributed subtest score, were computed as measures of item discrimination. Items with discrimination values above .80 are typically viewed as having high discriminatory power; items with discrimination values below .20 are typically viewed as having poor discriminatory power.

Computational formulas for these statistics differ according to whether the subtest is analyzed as a speeded or a power subtest. For a power subtest, item difficulty is calculated using all examinees taking the test, under the assumption that all examinees will have an opportunity to consider every subtest item. For a speeded subtest, difficulty is calculated using only examinees who respond to the item or a subsequent item of the subtest. Examinees who do not attempt items are not considered in these speeded analyses.

Subtest and Composite Analysis

Means, standard deviations, skew, kurtosis, reliability and proportion correct are presented for each subtest. For composite analyses, means, standard deviations, skew and kurtosis

values were calculated. Intercorrelation matrices are provided for the subtests and for the composites.

In addition to these subtest analyses using all test items, subtest statistics were calculated using only the common items for each subtest. These analyses check on the assumption of randomly equivalent samples.

Equating Analysis

Equating enables two forms of a test that are intended to be parallel, which are never precisely equivalent in level and range of difficulty, to be rendered interchangeable by converting the score units of one test to the score units of another. Statistical equating methods establish a relationship between raw scores on two test forms so that the score on one form can be used to express the score on the other form. In the current study, composite scores of Forms Q1 and Q2 were linked to the normative group using linear and equipercentile equating to Forms P scores (see Angoff, 1971 for further explanation of equating).

In linear equating, two raw scores are equated if their z-score values are equivalent, resulting in a smooth straight line. In equipercentile equating, two raw scores are equated if their percentile ranks are equivalent. Because equipercentile equating may result in irregular equating curves, three types of polynomial smoothing (linear, quadratic and cubic) are used, resulting in four possible equatings. The linear and equipercentile equating methods coincide when the score distributions are the same. In choosing from among the four possible equatings, the z-score linear equating and three polynomial smoothings, the sample descriptive statistics and size are among the characteristics to be considered. When the means, standard deviations, skew, and kurtosis of the two randomly equivalent equating samples are nearly identical on both tests being equated, the

z-score linear equating is to be preferred. Linear equating uses two parameters, the mean and standard deviation, per test form. When the z-score linear equating is not appropriate, then one of the three smoothings of equipercentile equatings is chosen. These polynomial smoothings are based upon two parameters for the linear smoothing, three parameters for the quadratic and four parameters for the cubic smoothings. The cubic smoothing of the polynomial equating fits the raw equipercentile data more closely than the quadratic, which fits more closely than the linear. When sample sizes and the range of scores on a test are large, the parameters of the cubic equating are stable and thus, cubic smoothed equipercentile equating should be considered.

Results and Discussion

Item Difficulty Analysis Results

For purposes of summation, item difficulty values are presented in a frequency distribution with five categories. The categories used are arbitrary and use of alternative categories would have changed the distribution. These categories have been used for similar purposes in reports of previous tests and are therefore retained for this report. The nature of the categories allow item difficulties within one category to be further apart than item difficulties between two categories. For example, an item difficulty of .41 is in the same category as a difficulty of .59, yet a different category than a difficulty of .39.

Because item difficulties are sample specific, distributions of item difficulties of the subtests are provided for each of the three subsamples as well as the total sample. However, we will focus our discussion on the distribution of item difficulties for the total sample presented in Table 5. Table 6 provides the summary statistics (mean, median, minimum and maximum) for the

item difficulty values for the total sample. Item difficulties of the subsamples across test forms are presented in Appendix A.

As is evident from Table 5, the majority of items in P1 have difficulties ranging from .20 to .80. Electrical Maze is the only subtest that includes items with difficulties below .20. Thirteen of the subtests have at least one item with a difficulty above .80. Approximately half of the items in the Table Reading subtest have item difficulties above .80, suggesting that Table Reading is a relatively easy subtest. Table 6 shows that all sixteen subtests have mean item difficulties between .40 and .60.

Form Q1 subtests have similar item difficulty characteristics as subtests in Form P1. Again, item difficulties tend to range from .20 to .80. Two subtests, Electrical Maze and Table Reading have items with item difficulties below .20. Thirteen subtests have at least one item with a difficulty value above .80. Table Reading is a relatively easy subtest; half of the items have difficulty values above .80. Fifteen subtests have a mean level of item difficulty between .40 and .60.

Item difficulties for test Form Q2 are predominantly in the .20 to .80 range. Three subtests, Verbal Analogies, Mechanical Comprehension and Electrical Maze, include items with item difficulties below .20. Twelve subtests include items with difficulty value greater than .80. As in P1 and Q2, the majority of items form the Table Reading subtests have difficulties above .80. Fifteen subtests have mean levels of item difficulty between .40 and .60.

Table 5. Distribution of Item Difficulties for Preliminary Equating Study

Subtest	P1					Q1					Q2				
	.00-.20	>.20-.40	>.40-.60	>.60-.80	>.80-.99	.00-.20	>.20-.40	>.40-.60	>.60-.80	>.80-.99	.00-.20	>.20-.40	>.40-.60	>.60-.80	>.80-.99
Verbal Analogies	0	2	6	10	7	0	1	7	8	9	1	1	4	11	8
Arithmetic Reasoning	0	1	9	10	5	0	3	8	11	3	0	1	7	11	6
Reading Comprehension	0	2	10	11	2	0	1	10	11	3	0	0	9	12	4
Data Interpretation	0	1	7	12	5	0	1	7	14	3	0	0	9	14	2
Word Knowledge	0	4	5	12	4	0	3	6	10	6	0	4	5	12	4
Math Knowledge	0	0	3	18	4	0	0	7	12	6	0	0	3	16	6
Mechanical Comprehension	0	6	11	3	0	0	3	13	4	0	1	5	9	5	0
Electrical Maze	2	10	5	3	0	2	9	5	4	0	2	8	4	6	0
Scale Reading	0	3	15	14	8	0	3	9	18	10	0	3	9	19	9
Instrument Comprehension	0	2	7	11	0	0	0	9	9	2	0	0	8	12	0
Block Counting	0	2	4	8	6	0	3	2	9	6	0	2	7	8	3
Table Reading	0	8	5	8	19	1	5	5	9	20	0	6	5	7	22
Aviation Information	0	5	10	3	1	0	3	11	6	0	0	5	8	6	1
Rotated Blocks	0	3	6	3	3	0	3	5	5	2	0	3	5	4	3
General Science	0	5	9	4	1	0	4	10	4	2	0	3	11	6	0
Hidden Figures	0	0	6	4	5	0	2	5	4	4	0	0	6	4	5

Table 6. Summary Statistics of Item Difficulties for Preliminary Equating Study

	P1				Q1				Q2			
Subtest	Mean	Median	Min	Max	Mean	Median	Min	Max	Mean	Median	Min	Max
Verbal Analogies	.587	.690	.208	.976	.599	.704	.353	.961	.593	.713	.184	.948
Arithmetic Reasoning	.563	.638	.335	.890	.551	.614	.321	.864	.574	.686	.369	.867
Reading Comprehension	.546	.602	.297	.842	.560	.629	.363	.874	.571	.660	.440	.852
Data Interpretation	.570	.642	.324	.913	.561	.652	.229	.853	.566	.632	.476	.882
Word Knowledge	.559	.666	.378	.923	.564	.660	.344	.894	.550	.675	.231	.898
Math Knowledge	.588	.690	.547	.886	.592	.710	.515	.893	.596	.725	.534	.879
Mechanical Comprehension	.491	.456	.295	.739	.501	.514	.222	.745	.494	.440	.155	.744
Electrical Maze	.451	.348	.160	.667	.454	.349	.132	.698	.466	.377	.097	.734
Scale Reading	.563	.647	.322	.910	.586	.698	.383	.912	.584	.708	.314	.947
Instrument Comprehension	.545	.622	.383	.755	.556	.657	.411	.819	.544	.632	.401	.746
Block Counting	.585	.677	.356	.928	.583	.693	.346	.933	.567	.623	.298	.928
Table Reading	.598	.795	.208	.934	.601	.793	.194	.956	.606	.827	.231	.933
Aviation Information	.497	.494	.236	.856	.504	.494	.256	.750	.509	.518	.282	.816
Rotated Blocks	.537	.545	.278	.891	.548	.600	.282	.884	.538	.577	.264	.890
General Science	.512	.520	.262	.882	.516	.469	.295	.880	.516	.525	.206	.783
Hidden Figures	.595	.672	.457	.934	.569	.640	.326	.916	.586	.649	.475	.921

There are fluctuations in the frequency distributions of the item difficulties on Forms P1, Q1 and Q2. When easier items are defined as those with difficulties greater than .60 then Q2 had two or more easier items than Q1 in Verbal Analogies, Arithmetic Reasoning, Reading Comprehension and Math Knowledge and Q1 had two or more easier items than Q2 in Block Counting. There are no substantial differences in the mean item difficulty of a subtest across the three test forms. The maximum difference in subtest mean item difficulty values among any two of the three test forms ranged from .004 to .026. Only four subtests, Arithmetic Reasoning, Reading Comprehension, Scale reading and Hidden Figures, had a largest pairwise difference greater than .020.

Item Discrimination Analysis Results

For purposes of summation, item discrimination values are presented in a frequency distribution with five categories. As is the case with the item difficulty distributions, the categories used are arbitrary and use of alternative categories would alter the distribution. These categories have been used for similar purposes in reports of previous tests and are therefore retained for this report. The nature of the categories allow item discriminations within one category to be further apart than item discrimination between two categories. For example, an item discrimination of .41 is in the same category as a discrimination of .59, yet a different category than a discrimination of .39.

Because item discriminations are sample specific, distributions of item discriminations are provided for each of the three subsamples as well as the total sample. However, we will focus our discussion on the frequency distribution of the total sample presented in Table 7. Readers who wish to compare item discriminations of the subsamples across test forms should refer to Appendix A. Table 8 provides the summary statistics for the item discrimination values for the total sample.

Table 7. Distribution of Biserial Correlations for Preliminary Equating Study

Subtest	P1					Q1					Q2				
	.00-.20	>.20 -.40	>.40 -.60	>.60 -.80	>.80 -.99	.00-.20	>.20 -.40	>.40 -.60	>.60 -.80	>.80 -.99	.00-.20	>.20 -.40	>.40 -.60	>.60 -.80	>.80 -.99
Verbal Analogies	0	0	9	13	3	0	1	9	13	2	0	0	8	14	3
Arithmetic Reasoning	0	0	3	14	8	0	0	2	20	3	0	0	2	18	5
Reading Comprehension	0	1	6	13	5	0	1	6	16	2	0	0	6	15	4
Data Interpretation	0	2	8	15	0	0	0	5	17	3	0	0	5	15	5
Word Knowledge	0	1	3	15	6	1	0	3	16	5	0	1	8	14	2
Math Knowledge	0	0	1	7	17	0	0	1	9	15	0	0	0	6	19
Mechanical Comprehension	0	3	8	9	0	0	0	10	10	0	0	3	6	10	1
Electrical Maze	0	1	13	6	0	0	5	8	7	0	0	4	6	10	0
Scale Reading	0	6	15	18	1	0	1	12	24	3	0	0	15	20	5
Instrument Comprehension	0	0	3	10	7	0	0	0	11	9	0	0	1	10	9
Block Counting	0	0	3	13	4	0	1	6	12	1	0	0	10	9	1
Table Reading	0	1	7	17	15	0	2	7	16	15	0	2	7	17	14
Aviation Information	0	0	9	8	2	0	0	8	10	2	0	1	5	10	4
Rotated Blocks	0	0	6	9	0	0	0	4	11	0	0	0	3	12	0
General Science	0	3	3	12	1	0	0	10	9	1	0	1	8	11	0
Hidden Figures	0	0	1	11	3	0	0	2	13	0	0	0	1	10	4

Table 8. Summary Statistics of Biserial Correlations for Preliminary Equating Study

	P1				Q1				Q2			
Subtest	Mean	Median	Min	Max	Mean	Median	Min	Max	Mean	Median	Min	Max
Verbal Analogies	.644	.622	.404	.849	.634	.653	.378	.827	.648	.623	.409	.896
Arithmetic Reasoning	.733	.741	.484	.879	.710	.723	.539	.817	.718	.769	.527	.881
Reading Comprehension	.668	.686	.350	.826	.648	.657	.379	.851	.705	.757	.499	.870
Data Interpretation	.603	.640	.258	.780	.659	.637	.448	.831	.693	.711	.431	.867
Word Knowledge	.707	.711	.205	.910	.709	.740	.193	.937	.655	.687	.278	.900
Math Knowledge	.822	.828	.590	.987	.811	.808	.592	1.00	.848	.844	.676	1.000
Mechanical Comprehension	.565	.573	.343	.764	.598	.588	.442	.760	.592	.640	.317	.863
Electrical Maze	.546	.546	.397	.699	.516	.521	.288	.670	.536	.540	.279	.723
Scale Reading	.581	.590	.215	.818	.636	.641	.390	.839	.650	.639	.421	.927
Instrument Comprehension	.758	.759	.543	.965	.757	.748	.601	.937	.778	.783	.547	.982
Block Counting	.703	.679	.410	.905	.632	.656	.315	.817	.624	.589	.418	.803
Table Reading	.727	.742	.303	.915	.716	.729	.320	.917	.716	.751	.241	.983
Aviation Information	.634	.609	.428	.811	.648	.641	.461	.845	.674	.647	.365	.912
Rotated Blocks	.602	.658	.559	.788	.667	.664	.485	.798	.653	.667	.528	.758
General Science	.608	.629	.315	.854	.596	.588	.424	.852	.601	.607	.355	.762
Hidden Figures	.723	.718	.569	.842	.705	.702	.557	.798	.734	.743	.519	.866

The items on all three test forms, P1, Q1 and Q2, show acceptable biserial correlations. The frequency distribution of biserial correlations in Table 7 shows that almost all are above .40 and the majority fall in the .60 to .80 range. The subtest mean biserial correlations in Table 8 are generally between .50 and .70 with the minimum mean biserial correlation values of .546, .516. and .536 for Forms P1, Q1, and Q2 respectively. These biserial correlations indicate that the dichotomous item responses correlate well with the subtest score and discriminate well among the examinees.

In comparing the subtest discrimination indices of P1, Q1, and Q2 it is evident that there are fluctuations in the frequency distributions of the biserial correlations. When item discriminations of greater than .60 are taken to be good discrimination then Q1 had five more such items in Word Knowledge than Q2 and three more in Block Counting, while Q2 had three more than Q1 in Electrical Maze. The maximum difference in subtest mean biserial correlation values for any two of the three test forms, P1, Q1, and Q2, ranged from .011 to .090. Q1 and Q2 had reasonably higher mean discrimination than P1 on Data Interpretation, Rotated Blocks and Scale Reading and lower discrimination than P1 on Block Counting. In comparing Forms Q1 and Q2, a difference in the subtest mean biserial correlations range from .000 to .057.

Subtests Analysis Results

Descriptive statistics for the subtests are provided for the total sample as well as the AFA, ROTC and BMTS subsamples. Table 9 provides the summary statistics for the subtests for the total sample. Readers who wish to compare subtests statistics of the subsamples across test forms should refer to Appendix A.

Table 9. Descriptive Statistics of Subtests for Preliminary Equating Study

	Proportion correct			Mean			Standard deviation			Skew			Kurtosis			Reliability		
Subtest	P1	Q1	Q2	P1	Q1	Q2	P1	Q1	Q2	P1	Q1	Q2	P1	Q1	Q2	P1	Q1	Q2
Verbal Analogies	.705	.730	.718	16.83	17.39	17.72	4.90	4.66	4.72	-.63	-.72	-.88	-.38	-.10	.36	.842	.828	.833
Arithmetic Reasoning	.652	.625	.678	15.94	15.38	16.63	6.31	6.24	6.28	-.30	-.25	-.44	-1.07	-1.04	-.97	.907	.899	.908
Reading Comprehension	.613	.644	.669	15.11	15.87	16.51	5.89	5.72	6.13	-.40	-.48	-.59	-.77	-.68	-.70	.882	.870	.896
Data Interpretation	.668	.648	.659	16.37	15.98	16.28	5.19	5.70	6.12	-.43	-.52	-.56	-.73	-.70	-.81	.841	.873	.892
Word Knowledge	.642	.653	.623	15.71	15.98	15.26	6.10	6.00	5.54	-.46	-.55	-.44	-.87	-.67	-.68	.897	.894	.869
Math Knowledge	.706	.716	.723	17.45	17.62	17.89	6.91	6.62	6.94	-.64	.-73	-.78	-.97	-.80	-.73	.931	.926	.936
Mechanical Comprehension	.477	.503	.486	9.55	10.06	9.76	4.30	4.52	4.41	.17	-.02	.07	-.80	-1.02	-.88	.786	.812	.807
Electrical Maze	.381	.388	.417	7.81	7.99	8.56	3.79	3.64	3.75	.53	.30	.23	.09	-.27	-.37	.745	.721	.743
Scale Reading	.651	.704	.699	25.30	27.43	27.23	7.86	8.10	8.30	-.32	-.71	-.62	-.73	-.26	-.51			
Instrument Comprehension	.609	.635	.607	12.12	12.56	12.08	5.61	5.43	5.76	-.36	-.36	-.33	-1.10	-1.10	-1.19	.900	.896	.908
Block Counting	.700	.697	.662	13.53	13.51	12.86	4.46	4.05	4.20	-.62	-.65	-.37	-.19	-.10	-.53	.855	.812	.814
Table Reading	.729	.735	.745	27.65	28.00	28.40	8.07	7.92	8.01	-.47	-.68	-.58	-.29	.13	-.24			
Aviation Information	.492	.509	.523	9.32	10.18	10.42	4.49	4.87	5.03	.12	.17	.15	-.93	-.96	-1.10	.829	.848	.862
Rotated Blocks	.590	.617	.594	8.59	9.09	8.73	3.38	3.45	3.40	-.07	-.42	-.30	-.84	-.60	-.62	.782	.786	.777
General Science	.531	.540	.539	9.97	10.64	10.73	4.17	4.26	4.50	-.07	-.19	-.16	-.84	-.70	-.76	.799	.795	.814
Hidden Figures	.723	.666	.702	10.44	9.65	10.21	3.40	3.48	3.60	-.43	-.17	-.45	-.66	-.88	-.69	.814	.813	.831

In general, the descriptive statistics of the subtests are similar across test forms. Subtest mean scores generally differed by less than one unit. Exceptions to this pattern, or subtest differences greater than one unit were observed between Forms P1 and Q1 on Scale Reading, between Forms P1 and Q2 on Reading Comprehension, Scale Reading and Aviation Information and between Forms Q1 and Q2 on Arithmetic Reasoning. The negligible magnitude of these differences provide support for the parallelism of these measures.

The skew and kurtosis values for the subtests are quite similar across test forms. The majority of the subtests are negatively skewed and none have skew values less than -1.00 or greater than +1.00. Kurtosis values are similar across test forms with a few values around -1.00, a value which indicates a slightly flatter score distribution. Thus, the subtest score distributions are relatively symmetric and tend toward normality.

Kuder-Richardson 20 reliability estimates provide evidence of generally high internal consistency and are approximately equivalent across test forms. The majority of the reliability values are greater than .80, and the lowest estimate is .721. Reliability estimates are not appropriate for subtests scored as speeded tests and thus are not provided for the Scale Reading and Table Reading subtests.

The subtest intercorrelation matrix is presented for Forms P1, Q1 and Q2 in Table 10. The data for all three forms are presented in one table to facilitate comparisons of subtest correlations across forms. The subtest intercorrelations are among the subtests within one form, not among subtests of different forms. The maximum correlation among subtests is .83, and represents the correlation between Arithmetic Reasoning and Data Interpretation subtests on Form Q2. The minimum correlation is .33 and occurs between the Word Knowledge and Electrical Maze

subtests on Form P1 and the Block Counting and Aviation Information subtests on Form Q1. The subtest intercorrelations show similar patterns across the three forms. The maximal difference between any of the three subtest correlations in the 120 triads is greater than .10 in only four cases; in these instances the correlation is either .10 or .11. Thus, there is a high degree of similarity among the correlation matrices across the three test forms.

The means of the common items on the subtests for Forms Q1 and Q2 are generally similar to one another and to those of Form P1. Table 1 in Appendix B presents the common item subtest means and standard deviations for forms P1, Q1 and Q2. The subtest means on Forms Q1 and Q2 tend to be slightly higher than the corresponding subtest means of Form P1, however twenty-seven out of thirty-two means are within one-tenth of a standard deviation of the P1 subtest means. The discrepancies occur for one of the comparisons between P1 and Q1 and for four of the comparisons between P1 and Q2. In general, the common item means across test forms are approximately equivalent.

Composite Analysis Results

The composite level statistics for Forms P1, Q1 and Q2 are reported for the total sample in Table 11. As would be expected given the similarity in the subtest characteristics, the composite scores are similar across test forms. Composite means for Forms Q1 and Q2 are generally closer than means of P1 with either of the Forms Q. The composite mean scores suggest that Forms Q1 and Q2 are slightly easier than P1, except for the Quantitative composite. Form Q2 has higher mean composite scores than Q1 on the Navigator-Technical, Academic Aptitude, and Quantitative composites while Form Q1 has higher mean composite scores on the Pilot and Verbal composites. However, there should be no significant differences in mean composite scores for Forms Q after the equating.

Table 10. Intercorrelations of Subtests for Preliminary Equating Study

Subtest		AR	RC	DI	WK	MK	MC	EM	SR	IC	BC	TR	AI	RB	GS	HF
VA	P1	.70	.76	.68	.76	.72	.57	.42	.60	.59	.52	.46	.57	.53	.69	.55
	Q1	.64	.78	.67	.77	.67	.57	.40	.58	.52	.44	.43	.46	.46	.66	.48
	Q2	.65	.75	.67	.74	.67	.54	.40	.57	.52	.46	.43	.53	.48	.65	.52
AR	P1		.67	.80	.63	.82	.61	.44	.75	.55	.56	.58	.58	.58	.68	.51
	Q1		.67	.80	.62	.78	.64	.45	.72	.55	.53	.55	.51	.54	.69	.48
	Q2		.71	.83	.60	.81	.59	.50	.78	.56	.58	.60	.53	.58	.68	.52
RC	P1			.68	.78	.69	.51	.38	.59	.54	.46	.49	.54	.42	.67	.49
	Q1			.68	.79	.68	.57	.39	.58	.50	.44	.45	.50	.46	.69	.45
	Q2			.77	.78	.71	.58	.44	.62	.53	.48	.49	.59	.49	.71	.47
DI	P1				.60	.77	.56	.45	.71	.56	.56	.60	.55	.52	.61	.49
	Q1				.62	.77	.64	.44	.73	.56	.56	.58	.54	.58	.67	.50
	Q2				.67	.82	.61	.51	.76	.57	.57	.59	.59	.59	.72	.54
WK	P1					.68	.48	.33	.52	.50	.42	.41	.57	.41	.66	.46
	Q1					.63	.58	.35	.52	.50	.40	.41	.53	.44	.69	.42
	Q2					.64	.54	.37	.50	.47	.40	.39	.56	.42	.68	.43
MK	P1						.55	.41	.71	.59	.55	.59	.57	.54	.69	.53
	Q1						.63	.44	.72	.58	.52	.57	.52	.55	.73	.52
	Q2						.55	.46	.71	.57	.54	.60	.58	.58	.70	.56
MC	P1							.47	.51	.58	.48	.41	.63	.60	.67	.51
	Q1							.47	.57	.64	.46	.38	.62	.62	.69	.51
	Q2							.49	.51	.60	.47	.36	.63	.54	.69	.47
EM	P1								.49	.52	.52	.42	.38	.46	.44	.46
	Q1								.46	.48	.43	.41	.36	.41	.42	.42
	Q2								.52	.48	.53	.39	.38	.45	.46	.45
SR	P1									.59	.64	.64	.53	.56	.56	.51
	Q1									.59	.62	.63	.50	.55	.61	.50
	Q2									.56	.61	.61	.48	.55	.59	.52
IC	P1										.55	.47	.61	.56	.60	.54
	Q1										.51	.46	.62	.58	.59	.56
	Q2										.55	.45	.59	.56	.59	.57
BC	P1											.58	.42	.55	.47	.55
	Q1											.57	.33	.53	.45	.50
	Q2											.57	.39	.52	.47	.53
TR	P1												.42	.44	.40	.43
	Q1												.37	.44	.44	.42
	Q2												.41	.49	.45	.44
AI	P1													.50	.62	.43
	Q1													.45	.59	.40
	Q2													.49	.66	.42
RB	P1														.54	.56
	Q1														.52	.58
	Q2														.56	.59
GS	P1															.54
	Q1															.49
	Q2															.53

Table 11. Descriptive Statistics of Composites for Preliminary Equating Study

	Mean			Standard deviation			Skew			Kurtosis		
Composite	P1	Q1	Q2	P1	Q1	Q2	P1	Q1	Q2	P1	Q1	Q2
Pilot	122.11	127.12	126.58	33.55	32.84	33.60	-.42	-.62	-.47	-.55	-.46	-.63
Navigator-Technical	162.61	165.34	167.28	45.63	45.74	47.42	-.38	-.59	-.53	-.69	-.60	-.71
Academic Aptitude	97.41	98.22	99.84	30.95	30.44	31.46	-.58	-.62	-.70	-.77	-.68	-.55
Verbal	47.64	49.24	49.03	15.57	15.17	15.06	-.54	-.60	-.68	-.70	-.54	-.36
Quantitative	49.77	48.97	50.80	17.15	17.20	18.15	-.47	-.51	-.60	-.97	-.95	-.89

The skew and kurtosis values for the composites are quite similar across the three test forms. The skew values range from -.38 to -.70; kurtosis values range from -.36 to -.97. These skew and kurtosis values indicate the composite score distributions are relatively symmetric and tend toward normality.

The composite intercorrelation matrix for Forms P1, Q1 and Q2 is presented in Table 12. The data for all three forms are presented in one table to facilitate comparisons of composite correlations across forms. The composite intercorrelations are among the composite within one form, not among composites of different forms. The maximum correlation among composites is .96 and results from the correlation between the Pilot and Navigator-Technical composites on all three forms. The minimum correlation is .75 and occurs between the Verbal and Pilot composites and Verbal and Navigator-Technical composites on Form Q1. The composite intercorrelations are almost identical across test forms; the maximum difference between any of the three composite correlations in a triads is .03. Thus, there is a high degree of similarity among the composite intercorrelation matrices across the three test forms.

Table 12. Intercorrelations of Composites for Preliminary Equating Study

Composite	Test form	Nav-tech	Academic Aptitude	Verbal	Quantitative
Pilot	P1	.96	.85	.76	.84
	Q1	.96	.85	.75	.85
	Q2	.96	.86	.77	.86
Navigator-Technical	P1		.90	.77	.93
	Q1		.90	.75	.93
	Q2		.91	.77	.94
Academic Aptitude	P1			.94	.95
	Q1			.93	.94
	Q2			.93	.95
Verbal	P1				.78
	Q1				.76
	Q2				.79

Equating Analysis Results

Four possible equatings, the z-score linear, linear smoothed equipercentile, quadratic smoothed equipercentile and cubic smoothed equipercentile, were developed and compared for each composite on Q1 and Q2. The lack of nearly identical moments (skew and kurtosis) for the score distributions ruled out the z-score linear equating method and given that sample sizes were large enough to ensure stability, the cubic smoothing equipercentile equatings were selected for each of the five composites on each test form. Using this equipercentile equating with cubic smoothing, preliminary conversion tables were developed and are presented in Appendix C.

OPERATIONAL EQUATING STUDY

Subjects

Subject samples for the operational equating study were actual examinees taking the AFOQT Forms P1, Q1, and Q2 for purposes of officer selection decisions, either into ROTC or into Air Force commissioning for those with college degrees. Their operational scores were provided by the preliminary conversion tables. These examinees were tested over a period from September of 1994 through June of 1995. On July 1, 1995, Forms Q1 and Q2 were pulled from the field while new equatings were accomplished using applicant scores.

Demographic information is presented for the total sample in Table 13. Based on the most frequent response within a demographic categorization, subjects were predominately male, Caucasian, with twelve or sixteen years of education and a high school degree or Bachelor's degree as the highest educational credential earned.

Table 13. Demographic Percentages of Total Sample for Operational Equating Study

		P1 n=4697	Q1 n=3387	Q n=300
Gender	Male	73.4	74.8	76.1
	Female	26.3	25.0	23.7
	Missing	0.3	0.3	0.2
Race	American Indian	0.8	0.6	0.9
	Asian	5.1	5.4	4.5
	Black	12.9	13.4	12.8
	Hispanic	6.7	6.1	6.6
	Caucasian	74.1	74.2	74.7
	Missing	0.3	0.4	0.4
Years of education	12	24.0	27.4	30.0
	13	13.6	14.5	14.7
	14	12.4	11.4	11.1
	15	13.6	14.4	13.0
	16	27.3	24.2	23.4
	17	5.2	4.3	4.4
	18	2.2	2.5	2.2
	19	0.7	0.5	0.7
	20	0.3	0.1	0.1
	21	0.0	0.2	0.1
	Missing	0.6	0.5	0.3
Highest degree earned	High School	58.3	63.0	65.0
	Associates	9.6	8.2	7.7
	Bachelors	29.8	26.5	24.7
	Masters	1.5	1.5	1.6
	Doctorate	0.2	0.2	0.2
	Missing	0.6	0.6	0.7

Administration

The AFOQT data for the operational equating study were collected from operational testing sessions at the Military Processing Stations (MEPS) and their outlying sites, Mobile Examining Team Sites (METS). Examiners followed the usual testing procedures for applicants,

with the exception that they were to cycle through Forms P1, Q1 and Q2 in that order to all examinees as they came in for testing.

Data Analysis

As mentioned previously, the data analysis section for both the preliminary and operational equating studies are similar. The main difference in the two analysis procedures and resultant output is that the preliminary analysis was comprised of total and subsample analyses, whereas the operational analyses involved no subgroup analyses. In addition, the second set of equating analyses, the operational equatings, allowed for comparisons between the preliminary and operational equatings based on the evaluation of critical selection cut areas.

Results and Discussion

Item Difficulty Analysis Results

As in the Preliminary Equating section, item difficulty values are presented in a frequency distribution with five categories. Distributions of item difficulties are provided in Table 14. Table 15 presents the summary statistics for the item difficulty values for the sample.

Table 14 shows the majority of items in P1 have difficulties ranging from .20 to .80. Electrical Maze and Table Reading are the only subtests that include items with difficulties below .20. Thirteen of the subtests have at least one item with a difficulty above .80. One-half of the items in the Table Reading subtest have item difficulties above .80, suggesting that Table Reading is a relatively easy subtest. The mean level of item difficulty for the subtests, shown in Table 15, is between .40 and .60 for all sixteen subtests.

Table 14. Distribution of Item Difficulties for Operational Equating Study

Subtest	P1					Q1					Q2				
	.00-.20	>.20-.40	>.40-.60	>.60-.80	>.80-.99	.00-.20	>.20-.40	>.40-.60	>.60-.80	>.80-.99	.00-.20	>.20-.40	>.40-.60	>.60-.80	>.80-.99
Verbal Analogies	0	4	7	8	6	0	3	8	6	8	1	1	7	9	7
Arithmetic Reasoning	0	6	7	8	4	0	6	7	11	1	0	4	6	10	5
Reading Comprehension	0	6	5	11	3	0	2	9	11	3	0	0	8	12	5
Data Interpretation	0	1	11	9	4	0	2	7	13	3	0	1	9	13	2
Word Knowledge	0	5	8	10	2	0	3	9	9	4	0	5	7	10	3
Math Knowledge	0	0	8	14	3	0	1	9	9	6	0	1	5	12	7
Mechanical Comprehension	0	10	7	3	0	1	7	9	3	0	1	9	7	3	0
Electrical Maze	6	6	6	2	0	4	7	6	3	0	2	9	3	6	0
Scale Reading	0	9	12	12	7	0	5	10	17	8	0	7	7	20	6
Instrument Comprehension	0	3	11	6	0	0	3	11	6	0	0	4	11	5	0
Block Counting	0	3	6	6	5	0	4	4	8	4	0	3	8	7	2
Table Reading	2	7	5	6	20	2	5	7	6	20	1	6	7	5	21
Aviation Information	0	11	6	1	1	1	11	5	3	0	1	8	8	3	0
Rotated Blocks	0	5	5	2	3	0	3	6	5	1	0	5	4	4	2
General Science	0	7	7	4	1	1	6	7	4	2	1	5	9	5	0
Hidden Figures	0	1	6	4	4	0	2	6	4	3	0	1	7	2	5

Table 15. Summary Statistics of Item Difficulties for Operational Equating Study

	P1				Q1				Q2			
Subtest	Mean	Median	Min	Max	Mean	Median	Min	Max	Mean	Median	Min	Max
Verbal Analogies	.567	.618	.201	.964	.582	.675	.274	.967	.583	.685	.179	.958
Arithmetic Reasoning	.545	.595	.309	.891	.535	.586	.267	.870	.558	.667	.279	.877
Reading Comprehension	.541	.605	.276	.828	.552	.632	.329	.867	.569	.643	.432	.875
Data Interpretation	.562	.609	.290	.908	.555	.626	.220	.883	.561	.653	.385	.924
Word Knowledge	.548	.595	.291	.915	.553	.604	.266	.899	.543	.604	.255	.912
Math Knowledge	.572	.655	.433	.903	.577	.693	.398	.929	.587	.693	.396	.895
Mechanical Comprehension	.475	.398	.273	.660	.476	.447	.172	.669	.477	.388	.120	.704
Electrical Maze	.438	.345	.120	.646	.439	.310	.123	.647	.451	.343	.088	.675
Scale Reading	.540	.571	.230	.909	.559	.644	.256	.904	.568	.660	.269	.959
Instrument Comprehension	.515	.553	.339	.698	.517	.527	.349	.788	.511	.534	.334	.708
Block Counting	.563	.622	.305	.917	.562	.646	.277	.913	.550	.583	.276	.915
Table Reading	.594	.800	.167	.938	.594	.799	.167	.958	.603	.815	.197	.943
Aviation Information	.458	.371	.202	.875	.461	.388	.190	.669	.467	.406	.180	.790
Rotated Blocks	.512	.464	.253	.864	.527	.561	.245	.865	.518	.497	.255	.879
General Science	.487	.440	.271	.820	.496	.420	.198	.874	.499	.503	.193	.788
Hidden Figures	.577	.634	.383	.939	.543	.559	.247	.893	.564	.594	.399	.903

The item difficulty distributions of subtests Form Q1 are similar to the item difficulty distributions of Form P1. Again, item difficulties tend to range from .20 to .80. Five subtests, Mechanical Comprehension, Electrical Maze, Table Reading, Aviation Information and General Science have items with item difficulties below .20. Twelve subtests have at least one item with a difficulty value above .80. Table Reading is a relatively easy subtest; half of the items have difficulty values above .80. All sixteen subtests have a mean level of item difficulty between .40 and .60.

Item difficulties for test Form Q2 occur predominantly in the .20 to .80 range. Six subtests, Verbal Analogies, Mechanical Comprehension, Electrical Maze, Table Reading, Aviation Information and General Science, include items with item difficulties below .20. Eleven subtests include items with difficulty value greater than .80. As in P1 and Q2, the majority of items from the Table Reading subtests have difficulties above .80. Fifteen subtests had mean level of item difficulty between .40 and .60.

The subtest difficulties of P1, Q1, and Q2 show fluctuations in the frequency distributions of the item difficulties. Consideration of items with difficulty greater than .60 gave the same results as found in the preliminary equating data. Q2 had two or more easier items than Q1 for Verbal Analogies, Arithmetic Reasoning, Reading Comprehension and Math Knowledge and Q1 had two or more easier items than Q2 on Block counting. There do not appear to be any substantial or systematic differences in the mean item difficulty of a subtest across the three test forms. The maximum difference in subtest mean item difficulty among any two of the three test forms ranged from .002 to .034. Only four subtests, Arithmetic Reasoning, Reading Comprehension, Scale Reading and Hidden Figures, had a largest pairwise difference above .020.

Item Discrimination Analysis Results

As in the Preliminary Equating section, biserial correlations are presented in a frequency distribution with five categories. Distributions of biserial correlations are provided in Table 16. Table 17 presents the summary statistics for the biserial correlation values for the sample.

The items on all three test forms, P1, Q1 and Q2, show acceptable biserial correlations. The frequency distribution of biserial correlations in Table 7 shows that the majority of the item biserial correlations fall in the .40 to .80 range. The subtest mean biserial correlations in Table 8 are generally between .50 and .70 with the minimum mean biserial correlation values of .511, .490. and .523 for Forms P1, Q1, and Q2 respectively. These biserial correlations indicate that the dichotomous item responses correlate well with the subtest score and discriminate well among the examinees.

Comparisons of the subtest discrimination indices of P1, Q1, and Q2 show that there are fluctuations in the frequency distributions of the biserial correlations. Investigation of items with discrimination greater than .60 showed different results than those found in the preliminary equating data. Q1 had five more such items than Q2 on Scale Reading and Instrument Comprehension and three more on Work Knowledge and Table Reading. Q2 had three more good discriminating items than Q1 on Data Interpretation. The maximum difference in subtest mean biserial correlation values for any two of the three test forms, P1, Q1, and Q2, ranged from .016 to .068. Data Interpretation, Mechanical Comprehension, and Aviation Information had reasonably higher mean biserials on Q1 and Q2 than on P1 and lower mean biserials on Block Counting than on P1. These results are somewhat similar to those found in the preliminary equating data. In comparing Forms Q1 and Q2, the difference in subtest mean biserial correlations range from .001 to .046.

Table 16. Distribution of Biserial Correlations for Operational Equating Study

Subtest	P1					Q1					Q2				
	.00-.20	>.20-.40	>.40-.60	>.60-.80	>.80-.99	.00-.20	>.20-.40	>.40-.60	>.60-.80	>.80-.99	.00-.20	>.20-.40	>.40-.60	>.60-.80	>.80-.99
Verbal Analogies	0	0	19	6	0	0	0	16	9	0	0	0	16	9	0
Arithmetic Reasoning	0	0	2	23	0	0	0	5	20	0	0	0	6	19	0
Reading Comprehension	0	2	11	12	0	0	3	10	12	0	0	3	11	11	0
Data Interpretation	0	4	14	7	0	0	0	18	7	0	0	3	12	10	0
Word Knowledge	0	1	5	18	1	0	1	8	16	0	0	1	11	13	0
Math Knowledge	0	0	4	20	1	0	0	4	20	1	0	0	5	16	4
Mechanical Comprehension	0	2	12	6	0	0	1	8	11	0	0	1	9	9	1
Electrical Maze	0	3	13	4	0	0	6	9	5	0	0	4	9	7	0
Scale Reading	0	7	24	9	0	0	2	23	15	0	0	1	29	10	0
Instrument Comprehension	0	0	4	15	1	0	0	4	11	5	0	0	3	14	3
Block Counting	0	0	4	15	1	0	2	5	13	0	0	0	12	8	0
Table Reading	0	2	8	20	10	0	4	9	21	6	0	2	14	16	8
Aviation Information	0	2	7	10	0	0	0	9	9	2	0	0	8	9	3
Rotated Blocks	0	0	4	11	0	0	0	4	11	0	0	0	4	11	0
General Science	0	2	8	9	0	0	0	14	6	0	0	1	14	5	0
Hidden Figures	0	0	1	14	0	0	0	4	11	0	0	1	4	10	0

Table 17. Summary Statistics of Biserial Correlations for Operational Equating Study

	P1				Q1				Q2			
Subtest	Mean	Median	Min	Max	Mean	Median	Min	Max	Mean	Median	Min	Max
Verbal Analogies	.520	.554	.406	.675	.561	.551	.436	.697	.544	.564	.416	.748
Arithmetic Reasoning	.664	.665	.441	.800	.647	.653	.436	.780	.646	.662	.464	.787
Reading Comprehension	.585	.585	.329	.759	.574	.574	.327	.787	.568	.563	.346	.788
Data Interpretation	.515	.519	.270	.695	.566	.562	.402	.711	.563	.569	.354	.711
Word Knowledge	.649	.687	.311	.801	.641	.637	.315	.800	.595	.613	.268	.781
Math Knowledge	.689	.723	.412	.852	.674	.675	.534	.824	.693	.699	.549	.835
Mechanical Comprehension	.537	.541	.301	.702	.593	.603	.385	.719	.581	.597	.388	.810
Electrical Maze	.518	.528	.347	.678	.490	.532	.258	.662	.523	.565	.329	.661
Scale Reading	.511	.505	.257	.756	.558	.575	.330	.739	.546	.538	.345	.751
Instrument Comprehension	.684	.680	.531	.873	.705	.715	.527	.878	.692	.693	.481	.919
Block Counting	.672	.677	.495	.859	.624	.622	.348	.788	.604	.587	.450	.758
Table Reading	.681	.717	.260	.850	.636	.639	.285	.871	.640	.690	.323	.882
Aviation Information	.590	.605	.358	.794	.641	.606	.485	.831	.640	.627	.435	.913
Rotated Blocks	.648	.651	.567	.740	.644	.628	.558	.756	.630	.622	.527	.722
General Science	.554	.600	.272	.675	.548	.542	.406	.635	.538	.532	.396	.683
Hidden Figures	.674	.673	.586	.781	.643	.661	.508	.723	.643	.656	.397	.787

Subtests Analysis Results

Table 18 provides the summary statistics for the subtests for the total sample. In general, the descriptive statistics of the subtests are similar across test forms. Subtest mean scores generally differed by less than one unit. Exceptions to this pattern, or subtest differences greater than one unit were observed between Forms P1 and Q1 on Scale Reading, between Forms P1 and Q2 on Reading Comprehension, Scale Reading and General Science and between Forms Q1 and Q2 on Arithmetic Reasoning and Scale Reading. The negligible magnitude of these differences provide support for the parallelism of these measures.

The skew and kurtosis values for the subtests are quite similar across test forms. The majority of the subtests are negatively skewed and none have skew values less than -1.00 or greater than +1.00. Kurtosis values are similar across test forms with a few values around -1.00, a value which indicates a slightly flatter score distribution. Thus, the subtest score distributions are relatively symmetric and tend toward normality.

Kuder-Richardson 20 reliability estimates provide evidence of generally high internal consistency and are quite similar across test forms. The majority of the reliability values are greater than .80, and the lowest estimate is .685. In general, these reliability values are lower than those obtained in the preliminary equating study. Reliability estimates are not appropriate for subtests scored as speeded tests and thus are not provided for the Scale Reading and Table Reading subtests.

The subtest intercorrelation matrix is presented for Forms P1, Q1 and Q2 in Table 19. The data for all three forms are presented in one table to facilitate comparisons of subtest correlations across forms. Again, the subtest intercorrelations are among the subtests within one form, not

among subtests of different forms. The maximum correlation among subtests is .76, the correlation between Arithmetic Reasoning and Data Interpretation subtests on Form Q2. The minimum correlation is .20 and occurs between the Word Knowledge and Electrical Maze subtests on Form P1. The subtest intercorrelations show similar patterns across the three forms. The maximal difference between any of the three subtest correlations in the 120 triads is greater than .10 in only two cases; in these instances the correlations are .10 and .11. Thus, there is a high degree of similarity among the correlation matrices across the three test forms.

The analyses of the common items on the subtests for Forms Q1 and Q2 indicate that the means are generally similar to one another and to those of Form P1. Table 2 in Appendix B presents the common item subtest means and standard deviations for forms P1, Q1 and Q2. The subtest means on Forms Q1 and Q2 tend to be slightly higher than the corresponding subtest means of Form P1, however twenty-six out of thirty-two means are within one-tenth of a standard deviation of the P1 subtest means. The discrepancies occur for one of the comparisons between P1 and Q1 and for five of the comparisons between P1 and Q2. In general, the common item means across test forms are approximately equivalent.

Table 18. Descriptive Statistics of Subtests for Operational Equating Study

	Proportion correct			Mean			Standard deviation			Skew			Kurtosis			Reliability		
Subtest	P1	Q1	Q2	P1	Q1	Q2	P1	Q1	Q2	P1	Q1	Q2	P1	Q1	Q2	P1	Q1	Q2
Verbal Analogies	.660	.695	.696	15.76	16.49	16.71	4.34	4.28	4.20	-.30	-.32	-.51	-.37	-.36	-.03	.779	.781	.774
Arithmetic Reasoning	.609	.585	.640	14.87	14.43	15.64	5.73	5.73	5.47	-.08	.03	-.12	-.88	-.90	-.86	.877	.871	.865
Reading Comprehension	.601	.627	.667	14.79	15.45	16.35	5.10	5.11	4.91	-.16	-.19	-.30	-.66	-.73	-.63	.833	.827	.820
Data Interpretation	.649	.633	.648	15.84	15.57	15.91	4.48	4.94	4.94	-.25	-.27	-.29	-.41	-.62	-.66	.770	.818	.815
Word Knowledge	.617	.629	.606	15.11	15.39	14.84	5.63	5.46	5.08	-.15	-.21	-.11	-.89	-.78	-.74	.870	.863	.835
Math Knowledge	.672	.683	.705	16.53	16.68	17.28	5.97	5.61	5.66	-.36	-.31	-.48	-.90	-.85	-.72	.888	.877	.883
Mechanical Comprehension	.440	.440	.444	8.82	8.87	8.98	4.07	4.48	4.32	.33	.25	.31	-.60	-.91	-.69	.755	.810	.797
Electrical Maze	.351	.352	.381	7.33	7.33	7.92	3.55	3.41	3.60	.44	.35	.41	.03	-.15	.05	.714	.685	.723
Scale Reading	.599	.642	.662	23.29	24.99	25.64	6.92	7.45	7.19	-.15	-.32	-.31	-.23	-.33	-.37			
Instrument Comprehension	.536	.543	.527	10.71	10.80	10.52	5.27	5.32	5.30	.08	.09	.10	-1.13	-1.18	-1.10	.872	.880	.875
Block Counting	.653	.649	.622	12.70	12.62	12.14	4.44	4.09	4.17	-.48	-.40	-.32	-.38	-.20	-.40	.841	.807	.801
Table Reading	.720	.721	.738	27.21	27.26	27.85	7.19	7.01	6.93	-.56	-.45	-.51	.63	.47	.53			
Aviation Information	.398	.404	.420	7.57	8.17	8.47	4.01	4.66	4.59	.90	.85	.86	.21	-.01	.01	.786	.836	.831
Rotated Blocks	.530	.567	.543	7.77	8.39	8.01	3.32	3.40	3.29	-.03	-.21	-.11	-.67	-.73	-.69	.770	.774	.756
General Science	.467	.491	.498	8.87	9.71	9.95	3.86	3.87	3.98	.24	.19	.11	-.50	-.45	-.58	.753	.750	.751
Hidden Figures	.683	.605	.655	9.79	8.82	9.51	3.20	3.26	3.28	-.29	.01	-.15	-.43	-.57	-.69	.776	.768	.772

Table 19. Intercorrelations of Subtests for Operational Equating Study

Subtest		AR	RC	DI	WK	MK	MC	EM	SR	IC	BC	TR	AI	RB	GS	HF
VA	P1	.59	.68	.57	.65	.51	.49	.32	.44	.40	.40	.26	.36	.43	.55	.39
	Q1	.55	.71	.60	.72	.46	.51	.28	.43	.36	.36	.26	.32	.37	.53	.37
	Q2	.56	.68	.58	.68	.44	.46	.31	.42	.37	.37	.29	.32	.36	.52	.32
AR	P1		.58	.72	.47	.70	.54	.42	.66	.43	.50	.38	.37	.51	.55	.42
	Q1		.57	.75	.48	.69	.59	.41	.66	.44	.45	.40	.35	.47	.57	.41
	Q2		.56	.76	.45	.71	.54	.42	.67	.43	.51	.43	.35	.46	.58	.41
RC	P1			.59	.71	.48	.41	.28	.44	.34	.38	.31	.33	.34	.51	.32
	Q1			.64	.72	.47	.49	.29	.46	.35	.36	.32	.34	.30	.54	.34
	Q2			.62	.70	.45	.48	.31	.43	.37	.36	.32	.37	.30	.56	.29
DI	P1				.46	.58	.49	.40	.64	.43	.51	.44	.37	.47	.48	.42
	Q1				.53	.64	.58	.40	.65	.46	.49	.44	.38	.47	.56	.44
	Q2				.50	.62	.58	.44	.64	.46	.50	.44	.40	.48	.59	.41
WK	P1					.36	.38	.20	.32	.27	.27	.21	.33	.26	.49	.25
	Q1					.36	.48	.21	.37	.31	.27	.23	.36	.28	.54	.28
	Q2					.34	.46	.23	.32	.31	.26	.21	.35	.26	.54	.23
MK	P1						.44	.36	.55	.35	.43	.37	.26	.46	.53	.41
	Q1						.48	.36	.54	.37	.38	.34	.24	.43	.56	.41
	Q2						.42	.35	.51	.33	.41	.35	.25	.40	.55	.39
MC	P1							.47	.46	.53	.44	.27	.50	.56	.62	.43
	Q1							.50	.54	.61	.46	.33	.56	.59	.68	.49
	Q2							.48	.48	.57	.45	.29	.53	.55	.67	.43
EM	P1								.48	.48	.46	.34	.32	.45	.39	.40
	Q1								.44	.47	.43	.34	.32	.44	.40	.41
	Q2								.47	.46	.51	.34	.33	.43	.41	.40
SR	P1									.50	.60	.54	.36	.50	.43	.47
	Q1									.48	.57	.56	.37	.48	.48	.45
	Q2									.45	.58	.54	.36	.46	.47	.44
IC	P1										.50	.36	.52	.53	.45	.44
	Q1										.46	.37	.57	.55	.50	.45
	Q2										.50	.34	.55	.50	.50	.44
BC	P1											.52	.29	.53	.36	.48
	Q1											.55	.27	.52	.36	.46
	Q2											.50	.30	.50	.38	.49
TR	P1												.25	.33	.22	.36
	Q1												.27	.35	.30	.34
	Q2												.27	.31	.25	.32
AI	P1													.35	.46	.28
	Q1													.38	.48	.29
	Q2													.31	.49	.26
RB	P1														.50	.52
	Q1														.47	.54
	Q2														.45	.50
GS	P1															.38
	Q1															.42
	Q2															.39

Composite Analysis Results

The composite level statistics for Forms P1, Q1 and Q2 are reported for the total sample in Table 20. As would be expected given the similarity in the subtest characteristics, the composite scores are similar across test forms. Composite means for Forms Q1 and Q2 are generally closer than means of P1 with either Form Q1 or Q2. The composite mean scores suggest that Forms Q1 and Q2 are slightly easier than P1, except for the Quantitative composite. Form Q2 has higher mean composite scores than Q1 on all composites, however, there should be no significant differences in mean composite scores for Forms Q after the equating.

The skew and kurtosis values for the composites are quite similar across the three test forms. The skew values range from -.14 to -.28; kurtosis values range from -.10 to -.80. These skew and kurtosis values indicate the composite score distributions are relatively symmetric and tend toward normality.

The composite intercorrelation matrix for Forms P1, Q1 and Q2 is presented in Table 21. The data for all three forms are presented in one table to facilitate comparisons of composite correlations across forms. The composite intercorrelations are among the composite within one form, not among composites of different forms. The maximum correlation among composites is .93 and results from the correlation between the Pilot and Navigator-Technical composites on all three Forms. The minimum correlation is .60 and occurs between the Pilot and Verbal composites on Form P1. The composite intercorrelations are nearly identical across test forms; the maximum difference between any of the three composite correlations in a triads is .02. Thus, there is a high degree of similarity among the composite intercorrelation matrices across the three test forms.

Table 20. Descriptive Statistics of Composites for Operational Equating Study

	Mean			Standard deviation			Skew			Kurtosis		
Composite	P1	Q1	Q2	P1	Q1	Q2	P1	Q1	Q2	P1	Q1	Q2
Pilot	113.40	116.52	118.24	28.36	29.20	28.63	-.19	-.15	-.19	-.10	-.20	-.18
Navigator-Technical	153.02	154.66	158.82	38.49	39.32	38.66	-.27	-.15	-.27	-.15	-.35	-.30
Academic Aptitude	92.89	94.02	96.71	25.16	25.33	24.37	-.22	-.21	-.28	-.55	-.60	-.55
Verbal	45.65	47.34	47.89	13.41	13.40	12.65	-.17	-.21	-.26	-.68	-.67	-.51
Quantitative	47.23	46.67	48.82	14.32	14.54	14.36	-.20	-.14	-.25	-.72	-.80	-.75

Table 21. Intercorrelations of Composites for Operational Equating Study

Composite	Test Form	Nav-tech	Academic Aptitude	Verbal	Quantitative
Pilot	P1	.93	.74	.60	.74
	Q1	.93	.75	.62	.73
	Q2	.93	.75	.62	.73
Navigator Technical	P1		.83	.62	.89
	Q1		.84	.62	.88
	Q2		.84	.61	.89
Academic Aptitude	P1			.90	.91
	Q1			.89	.91
	Q2			.88	.91
Verbal	P1				.64
	Q1				.64
	Q2				.62

Equating Analysis Results

Four possible equatings, the z-score linear, linear smoothed equipercentile, quadratic smoothed equipercentile and cubic smoothed equipercentile, were developed and compared for each composite on Q1 and Q2. As was the case in the preliminary equating study, the evaluations of the equatings ruled out the z-score linear equating and given that sample sizes were large enough to ensure stability, the cubic smoothing equipercentile equatings were selected for each of the five composites on each test form. Using this equipercentile equating with cubic smoothing, preliminary conversion tables were developed and are presented in Appendix D.

IMPLEMENTATION EFFECTS OF INSTITUTING THE OPERATIONAL CONVERSION TABLES

The preliminary conversion tables were used during the selection and classification of officer commissioning applicants during the data collection for the operational equating study. The data from the operational equating study were used to develop the operational equating

tables, which were not identical to the preliminary conversion tables. Minor discrepancies in the conversion tables were expected due to the differences in the samples used for the preliminary and operational equatings and to the differences in external motivation for testing. The sample of officer commissioning applicants used in the operational equating was larger, took the test to get in the Air Force instead of experimentally, and took the test under the actual MEPS administration conditions instead of a large group administration, and thus equatings developed on this sample were preferable. However, it was important to determine if the introduction of the operational tables would cause significant changes in qualification rates for officer positions. Qualification is determined by minimum cut-off values on some or all AFOQT composites for occupational categories such as pilot, navigator, missile, technical and non-line officers depending on the commissioning source of AFROTC, OTS, or the Airmen Enlisted Commissioning Program (AECP).

To examine the effects of the operational conversion tables, the various minimum cut-off values for officer categories and commissioning sources were identified and the raw score conversions to percentiles for both the preliminary and operational conversion tables were listed for a range of percentiles about those minima. The two conversion tables were very close except for the Navigator-Technical composite on Form Q2 at the tenth percentile. ROTC pilot qualification requires a minimum percentile of 50 on the Pilot composite and a 10 on the Navigator-Technical composite for applicants without a pilot's license and requires a minimum percentile of 25 on the Pilot composite with a 10 on the Navigator-Technical composite for applicants with a pilot's license. A distribution of applicants in the operational equating sample with Pilot Composite scores of 50 through 59 (n=367) showed none with a Navigator-Technical

score anywhere as low as the tenth percentile. A distribution of applicants in the operational equating sample with Pilot composite scores of 25 through 34 (n=352) found only three cases with a Navigator-Technical percentile less than 10 and only 8 cases with a Navigator-Technical percentile less than 15. Therefore, the tenth percentile minimum is basically an irrelevant minimum, so there will be no noticeable operational effect in switching from the preliminary conversion tables to the operational conversion tables.

CONCLUSIONS AND RECOMMENDATIONS

The AFOQT Forms Q1 and Q2 operational conversion tables based on the operational equating study should be implemented for use in making officer selection decisions. The operational conversion tables are more acceptable than the preliminary conversion tables because they were based on the responses of the larger, more appropriate sample used in the operational equating study. In the operational equating study the subjects were actual applicants for officer commissioning who were motivated to do well, thus the operational conversions tables developed on this sample are preferable.

REFERENCES

Angoff, W. H. (1971). Scales, norms, and equivalent scores. In R. L. Thorndike (Ed.), Educational Measurement (2nd ed.). Washington, DC: American Council on Education.

Berger, F. R., Gupta, W. B., Berger, R. M., & Skinner, J. (1988). Air Force Officer Qualifying Test (AFOQT) Form P: Test Construction (AFHRL-TR-88-30, AD-A200 678). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Berger, F. R., Gupta, W. B., Berger, R. M., & Skinner, J. (1990). Air Force Officer Qualifying Test (AFOQT) Form P: Test Manual (AFHRL-TR-89-56, AD-A221 004). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Gulliksen, H. (1950). Theory of mental tests. New York: John Wiley & Sons, Inc.

Rogers, D. L., Roach, B. W., & Short, L. O. (1986). Air Force Officer Qualifying Test Form O: Development and Standardization (AFHRL-TR-86-24, AD-A172 037). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Steuck, K. W., Watson, T. W., & Skinner, J. (1988). Air Force Officer Qualifying Test (AFOQT): Forms P pre-implementation analyses and equating (AFHRL-TR-88-6, AD-A201 100). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

APPENDIX A: RESULTS OF ANALYSES FOR AFA, ROTC , AND BMTS SUBSAMPLES FOR PRELIMINARY EQUATING STUDY

Table A-1. Distribution of Item Difficulties for Form P1 for AFA, ROTC and BMTS Subsamples in Preliminary Equating Study

Subtest	AFA					ROTC					BMTS				
	.00-.20	>.20-.40	>.40-.60	>.60-.80	>.80-.99	.00-.20	>.20-.40	>.40-.60	>.60-.80	>.80-.99	.00-.20	>.20-.40	>.40-.60	>.60-.80	>.80-.99
Verbal Analogies	1	1	1	7	15	0	1	4	11	9	4	9	5	4	3
Arithmetic Reasoning	0	0	2	9	14	0	1	8	8	8	5	11	6	3	0
Reading Comprehension	0	2	4	8	11	0	2	6	7	10	8	7	8	2	0
Data Interpretation	0	0	2	15	8	0	1	5	13	6	1	12	7	5	0
Word Knowledge	0	0	5	6	14	0	0	8	9	8	5	11	7	1	1
Math Knowledge	0	0	0	2	23	0	0	1	14	10	3	15	6	1	0
Mechanical Comprehension	0	3	8	7	2	0	5	8	6	1	2	15	3	0	0
Electrical Maze	1	11	2	6	0	2	9	2	7	0	8	7	5	0	0
Scale Reading	0	0	12	9	19	0	1	14	14	11	6	15	8	10	1
Instrument Comprehension	0	0	4	11	5	0	0	4	11	5	2	14	4	0	0
Block Counting	0	0	4	9	7	0	1	5	7	7	3	5	4	6	2
Table Reading	0	4	5	6	25	0	6	5	5	24	11	5	5	4	15
Aviation Information	0	1	5	7	6	0	6	8	4	1	8	9	1	1	0
Rotated Blocks	0	3	2	7	3	0	3	3	5	4	3	6	3	3	0
General Science	1	1	8	4	5	0	3	7	6	3	3	12	3	1	0
Hidden Figures	0	0	3	6	6	0	0	2	6	7	1	6	3	3	2

Table A-2. Distribution of Item Difficulties for Form Q1 for AFA, ROTC, and BMTS Subsamples in Preliminary Equating Study

Subtest	AFA					ROTC					BMTS				
	.00-.20	>.20-.40	>.40-.60	>.60-.80	>.80-.99	.00-.20	>.20-.40	>.40-.60	>.60-.80	>.80-.99	.00-.20	>.20-.40	>.40-.60	>.60-.80	>.80-.99
Verbal Analogies	0	0	5	7	13	0	0	6	7	12	2	9	5	5	4
Arithmetic Reasoning	0	0	5	7	13	0	1	7	10	7	8	8	7	2	0
Reading Comprehension	0	0	4	12	9	0	0	7	11	7	5	9	9	2	0
Data Interpretation	0	1	1	10	13	0	1	5	14	5	3	10	9	3	0
Word Knowledge	0	0	5	4	16	0	2	5	10	8	4	12	7	1	1
Math Knowledge	0	0	0	4	21	0	0	2	9	14	5	11	6	3	0
Mechanical Comprehension	0	2	6	10	2	0	3	12	4	1	3	13	4	0	0
Electrical Maze	2	7	6	5	0	3	7	4	6	0	7	9	4	0	0
Scale Reading	0	0	3	11	26	0	0	9	12	19	4	14	13	9	0
Instrument Comprehension	0	0	5	7	8	0	0	4	10	6	3	12	5	0	0
Block Counting	0	0	5	9	6	0	0	5	5	10	3	6	3	5	3
Table Reading	0	4	3	7	26	0	6	4	5	25	11	5	3	10	11
Aviation Information	0	0	6	6	8	0	5	8	6	1	5	11	4	0	0
Rotated Blocks	0	1	2	7	5	0	1	4	7	3	1	7	5	2	0
General Science	0	1	7	6	6	0	3	10	5	2	6	10	3	1	0
Hidden Figures	0	1	5	3	6	0	0	3	6	6	2	7	2	3	1

Table A-3. Distribution of Item Difficulties for Form Q2 for AFA, ROTC and BMTS Subsamples in Preliminary Equating Study

Subtest	AFA					ROTC					BMTS				
	.00-.20	>.20-.40	>.40-.60	>.60-.80	>.80-.99	.00-.20	>.20-.40	>.40-.60	>.60-.80	>.80-.99	.00-.20	>.20-.40	>.40-.60	>.60-.80	>.80-.99
Verbal Analogies	1	1	2	6	15	0	1	4	9	11	1	9	7	5	3
Arithmetic Reasoning	0	0	2	4	19	0	0	7	10	8	3	8	8	6	0
Reading Comprehension	0	0	0	10	15	0	0	6	10	9	1	12	10	2	0
Data Interpretation	0	0	0	8	17	0	0	7	14	4	3	10	10	2	0
Word Knowledge	0	2	4	4	15	0	3	7	8	7	6	6	11	1	1
Math Knowledge	0	0	0	3	22	0	0	3	12	10	4	13	7	1	0
Mechanical Comprehension	1	2	6	7	4	1	4	11	3	1	4	10	6	0	0
Electrical Maze	1	8	4	7	0	2	8	3	7	0	8	6	6	0	0
Scale Reading	0	0	6	10	24	0	2	6	16	16	3	13	14	9	1
Instrument Comprehension	0	0	4	8	8	0	0	5	14	1	3	10	7	0	0
Block Counting	0	0	4	13	3	0	2	2	8	8	3	5	4	6	2
Table Reading	0	3	3	5	29	0	6	4	6	24	10	4	4	8	14
Aviation Information	0	1	3	5	11	0	6	7	6	1	6	11	3	0	0
Rotated Blocks	0	2	1	7	5	0	1	6	5	3	3	5	4	3	0
General Science	0	2	4	7	7	1	1	11	5	2	2	14	4	0	0
Hidden Figures	0	0	4	5	6	0	0	3	6	6	1	6	4	3	1

Table A-4. Distribution of Biserial Correlations for Form P1 for AFA, ROTC and BMTS Subsamples in Preliminary Equating Study

Subtest	AFA					ROTC					BMTS				
	.00-.20	>.20-.40	>.40-.60	>.60-.80	>.80-.99	.00-.20	>.20-.40	>.40-.60	>.60-.80	>.80-.99	.00-.20	>.20-.40	>.40-.60	>.60-.80	>.80-.99
Verbal Analogies	0	3	19	3	0	0	3	15	7	0	0	4	11	10	0
Arithmetic Reasoning	0	1	7	8	9	0	0	5	16	4	0	3	9	13	0
Reading Comprehension	1	4	14	6	0	0	5	8	10	2	0	3	9	13	0
Data Interpretation	2	4	15	4	0	1	2	13	9	0	2	6	15	2	0
Word Knowledge	1	4	12	6	2	0	1	8	15	1	0	1	11	12	1
Math Knowledge	0	3	9	10	3	0	0	4	17	4	0	1	15	9	0
Mechanical Comprehension	0	2	14	4	0	0	3	12	5	0	0	7	12	1	0
Electrical Maze	0	2	14	4	0	0	1	14	5	0	0	7	9	4	0
Scale Reading	5	12	14	8	1	1	9	14	16	0	0	13	24	3	0
Instrument Comprehension	0	0	5	13	2	0	0	3	10	7	0	4	5	9	2
Block Counting	0	1	8	9	2	0	2	8	7	3	0	0	3	13	4
Table Reading	0	2	9	13	16	0	1	8	18	13	0	2	11	16	11
Aviation Information	0	2	13	4	0	0	3	10	6	0	0	7	12	0	0
Rotated Blocks	0	0	8	7	0	0	0	5	10	0	0	1	6	8	0
General Science	0	5	9	5	0	0	2	9	8	0	0	5	10	4	0
Hidden Figures	0	2	5	8	0	0	0	1	12	2	0	0	5	9	1

Table A-5. Distribution of Biserial Correlations for Form Q1 for AFA, ROTC and BMTS Subsamples in Preliminary Equating Study

Subtest	AFA					ROTC					BMTS				
	.00-.20	>.20-.40	>.40-.60	>.60-.80	>.80-.99	.00-.20	>.20-.40	>.40-.60	>.60-.80	>.80-.99	.00-.20	>.20-.40	>.40-.60	>.60-.80	>.80-.99
Verbal Analogies	1	5	17	2	0	1	3	16	4	1	0	4	19	2	0
Arithmetic Reasoning	0	3	12	7	3	0	0	13	12	0	0	5	15	5	0
Reading Comprehension	0	5	12	6	2	0	5	11	9	0	0	3	16	6	0
Data Interpretation	0	6	14	5	0	0	0	17	8	0	0	4	15	6	0
Word Knowledge	0	5	9	10	1	1	0	8	15	1	0	0	10	15	0
Math Knowledge	2	4	9	8	2	0	0	2	20	3	0	5	15	5	0
Mechanical Comprehension	0	2	15	3	0	0	3	7	10	0	1	5	14	0	0
Electrical Maze	1	8	6	5	0	0	3	7	10	0	0	7	12	1	0
Scale Reading	2	11	14	12	1	0	6	22	12	0	2	8	23	7	0
Instrument Comprehension	0	0	6	11	3	0	0	3	10	7	0	2	11	6	1
Block Counting	0	7	4	9	0	0	4	6	8	2	0	1	7	11	1
Table Reading	1	4	13	13	9	0	2	9	20	9	2	1	9	19	9
Aviation Information	0	5	13	2	0	0	0	8	12	0	1	6	13	0	0
Rotated Blocks	0	0	7	7	1	0	0	7	7	1	0	1	5	9	0
General Science	1	6	11	2	0	0	3	17	0	0	0	4	13	3	0
Hidden Figures	0	1	5	8	1	0	0	1	13	1	0	1	10	4	0

Table A-6. Distribution of Biserial Correlations for Form Q2 for AFA, ROTC and BMTS Subsamples in Preliminary Equating Study

Subtest	AFA					ROTC					BMTS				
	.00-.20	>.20-.40	>.40-.60	>.60-.80	>.80-.99	.00-.20	>.20-.40	>.40-.60	>.60-.80	>.80-.99	.00-.20	>.20-.40	>.40-.60	>.60-.80	>.80-.99
Verbal Analogies	1	7	16	1	0	1	2	14	8	0	0	0	16	9	0
Arithmetic Reasoning	2	3	8	7	5	0	0	3	20	2	0	0	16	8	1
Reading Comprehension	0	6	12	6	1	0	1	11	10	3	0	2	13	8	2
Data Interpretation	0	8	14	3	0	0	2	7	15	1	0	1	17	7	0
Word Knowledge	1	1	19	4	0	0	2	13	9	1	0	2	9	14	0
Math Knowledge	0	8	5	9	3	0	0	5	11	9	0	0	14	10	1
Mechanical Comprehension	0	2	11	6	1	0	3	8	8	1	0	3	13	4	0
Electrical Maze	3	3	8	6	0	0	5	7	8	0	1	5	9	5	0
Scale Reading	2	14	13	9	2	0	3	16	17	4	0	7	23	10	0
Instrument Comprehension	0	0	3	10	7	0	0	4	10	6	0	1	7	10	2
Block Counting	0	3	12	5	0	0	3	7	9	1	0	1	7	11	1
Table Reading	3	5	13	10	9	0	1	8	19	12	0	5	11	13	11
Aviation Information	0	3	8	8	1	0	0	10	9	1	1	8	10	1	0
Rotated Blocks	0	1	7	7	0	0	0	5	10	0	0	1	8	6	0
General Science	0	8	12	0	0	0	2	11	7	0	0	2	15	3	0
Hidden Figures	1	1	2	10	1	0	0	2	10	3	0	1	4	10	0

Table A-7. Descriptive Statistics of Subtests for AFA Subsample in Preliminary Equating Study

	Proportion correct			Mean			Standard deviation			Skew			Kurtosis			Reliability		
Subtest	P1	Q1	Q2	P1	Q1	Q2	P1	Q1	Q2	P1	Q1	Q2	P1	Q1	Q2	P1	Q1	Q2
Verbal Analogies	.826	.840	.835	19.35	19.73	19.51	2.99	2.82	2.71	-.41	-.32	-.23	-.38	-.36	-.39	.634	.589	.567
Arithmetic Reasoning	.834	.808	.871	20.06	19.45	20.89	3.96	4.03	3.31	-.76	-.61	-.68	-.15	-.40	-.52	.809	.798	.760
Reading Comprehension	.746	.768	.835	18.07	18.74	20.32	3.78	3.94	3.50	-.37	-.75	-.92	-.54	.52	.25	.727	.751	.740
Data Interpretation	.806	.821	.855	19.50	19.88	20.80	3.23	3.11	2.73	-.67	-.89	-1.03	-.04	.99	1.58	.650	.654	.582
Word Knowledge	.825	8.16	.792	19.58	19.37	18.66	3.35	3.65	3.16	-.87	-.74	-.56	.30	.95	.06	.713	.758	.661
Math Knowledge	.915	.913	.943	22.52	22.25	23.11	2.44	2.39	1.85	-1.23	-1.43	-1.31	1.13	3.10	1.91	.663	.622	.537
Mechanical Comprehension	.591	.651	.627	11.65	12.78	12.30	3.75	3.46	3.70	.05	-.41	-.46	-.74	-.16	-.44	.725	.684	.741
Electrical Maze	.427	.426	.469	8.66	8.71	9.50	3.76	2.99	3.22	.60	.11	.14	.39	-.01	-.25	.731	.565	.634
Scale Reading	.770	.828	.831	29.48	31.95	31.87	5.34	5.28	5.07	-.52	-1.11	-.84	-.22	1.82	.68			
Instrument Comprehension	.708	.760	.735	13.96	14.81	14.46	4.30	3.97	4.51	-.51	-.73	-.83	-.79	-.16	-.20	.828	.819	.861
Block Counting	.780	.744	.723	15.00	14.33	14.05	3.61	3.34	3.30	-.60	-.58	-.23	-.26	-.19	-.33	.782	.719	.681
Table Reading	.841	.838	.869	31.63	31.53	32.75	6.22	5.44	5.15	-1.01	-.88	-.76	2.35	2.21	.87			
Aviation Information	.718	.744	.795	13.03	14.35	15.28	2.94	2.80	2.87	-.18	-.31	-.75	-.57	-.10	.39	.640	.580	.662
Rotated Blocks	.672	.722	.725	9.65	10.45	10.43	2.88	2.90	2.69	-.18	-.66	-.57	-.94	.16	.02	.701	.719	.667
General Science	.686	.737	.706	12.09	13.64	13.69	2.83	2.80	2.83	-.01	.04	-.21	-.59	-.50	-.01	.589	.562	.561
Hidden Figures	.789	.726	.757	11.25	10.29	10.91	2.65	2.96	3.05	-.40	-.28	-.39	-.49	-.61	-.88	.698	.743	.767

Table A-8. Descriptive Statistics of Subtests for ROTC Subsample in Preliminary Equating Study

	Proportion correct			Mean			Standard deviation			Skew			Kurtosis			Reliability		
Subtest	P1	Q1	Q2	P1	Q1	Q2	P1	Q1	Q2	P1	Q1	Q2	P1	Q1	Q2	P1	Q1	Q2
Verbal Analogies	.781	.805	.783	18.36	18.85	18.52	3.57	3.40	3.52	-.52	-.66	-.69	-.22	.53	.71	.713	.700	.711
Arithmetic Reasoning	.705	.683	.720	17.07	16.65	17.56	5.39	5.32	5.53	-.33	-.37	-.52	-.83	-.58	-.59	.871	.861	.881
Reading Comprehension	.693	.720	.727	16.81	17.57	17.79	4.73	4.42	4.99	-.34	-.47	-.65	-.67	-.21	-.30	.822	.786	.844
Data Interpretation	.716	.697	.701	17.46	17.09	17.24	4.40	4.67	5.05	-.53	-.57	-.64	-.17	-.09	-.25	.784	.809	.836
Word Knowledge	.700	.722	.670	16.96	17.41	16.24	5.00	4.75	4.69	-.37	-.55	-.33	-.71	-.42	-.50	.845	.832	.815
Math Knowledge	.784	.807	.792	19.21	19.65	19.44	5.16	4.70	5.19	-.81	-1.02	-.87	-.25	.65	-.19	.877	.860	.884
Mechanical Comprehension	.512	.540	.504	10.20	10.75	10.07	4.02	4.17	4.16	.05	-.14	.12	-.77	-.85	-.73	.755	.777	.782
Electrical Maze	.422	.438	.458	8.60	8.93	9.28	3.65	3.67	3.71	.43	.19	.18	-.16	-.34	-.28	.725	.733	.745
Scale Reading	.703	.768	.749	27.11	29.67	28.92	6.79	5.97	7.11	-.30	-.66	-.80	-.64	.06	.13			
Instrument Comprehension	.711	.732	.688	14.05	14.37	13.63	4.75	4.50	5.04	-.79	-.64	-.60	-.25	-.64	-.80	.870	.858	.881
Block Counting	.762	.762	.717	14.61	14.71	13.87	3.68	3.39	3.79	-.55	-.76	-.44	-.18	.37	-.57	.794	.749	.787
Table Reading	.758	.784	.766	28.80	29.77	29.30	7.48	6.84	7.61	-.56	-.64	-.69	-.04	.20	.14			
Aviation Information	.516	.523	.519	9.70	10.42	10.34	3.85	4.43	4.28	.30	.44	.41	-.65	-.52	-.45	.756	.815	.804
Rotated Blocks	.656	.678	.627	9.42	9.88	9.16	3.14	2.97	3.18	-.23	-.58	-.41	-.60	-.04	-.24	.753	.717	.745
General Science	.589	.589	.578	10.93	11.43	11.46	3.77	3.35	3.97	-.09	.05	-.13	-.70	-.65	-.44	.752	.655	.757
Hidden Figures	.796	.755	.796	11.50	10.85	11.46	3.00	3.13	2.99	-.71	-.48	-.70	-.16	-.41	-.20	.785	.791	.781

Table A-9. Descriptive Statistics of Subtests for BMTS Subsample in Preliminary Equating Study

	Proportion correct			Mean			Standard deviation			Skew			Kurtosis			Reliability		
Subtest	P1	Q1	Q2	P1	Q1	Q2	P1	Q1	Q2	P1	Q1	Q2	P1	Q1	Q2	P1	Q1	Q2
Verbal Analogies	.437	.478	.497	10.96	11.92	12.42	4.15	4.08	5.09	.58	.29	.15	.11	-.42	-.72	.752	.737	.833
Arithmetic Reasoning	.357	.325	.397	9.31	8.58	10.20	4.68	4.15	5.01	.99	1.16	.86	.73	1.72	.15	.807	.744	.821
Reading Comprehension	.319	.360	.389	8.93	9.34	9.91	4.66	4.55	5.26	.90	.77	.57	.78	.20	-.30	.807	.780	.838
Data Interpretation	.429	.369	.373	10.80	9.65	9.56	3.96	4.49	4.94	.62	.59	.85	.37	-.03	.45	.685	.772	.813
Word Knowledge	.352	.370	.377	9.01	9.42	9.68	5.05	5.39	5.03	.96	.75	.71	.70	-.09	.08	.839	.857	.832
Math Knowledge	.327	.327	.359	8.46	8.56	9.17	4.50	3.89	5.42	1.37	1.00	.93	2.28	1.52	.43	.781	.702	.851
Mechanical Comprehension	.294	.280	.311	6.01	5.79	6.44	3.10	2.97	3.44	.91	.88	.72	1.12	.90	.43	.612	.589	.696
Electrical Maze	.250	.255	.288	5.29	5.35	6.11	2.94	2.73	3.29	.78	.81	.82	1.05	1.64	.65	.616	.538	.680
Scale Reading	.419	.445	.468	17.19	18.08	18.85	6.07	6.85	7.34	.47	.27	.42	.39	-.40	-.23			
Instrument Comprehension	.304	.306	.314	6.20	6.35	6.48	4.11	3.64	4.50	.99	.91	.93	.62	.82	.29	.797	.735	.839
Block Counting	.475	.511	.473	9.73	10.19	9.58	4.57	4.06	4.11	-.05	.05	.21	-.70	-.45	-.46	.862	.803	.811
Table Reading	.506	.503	.549	20.79	20.69	21.94	6.96	7.51	7.30	.08	-.13	.07	.01	-.21	-.11			
Aviation Information	.222	.257	.248	4.52	5.40	5.35	2.34	2.74	2.55	.67	1.13	.79	1.10	1.91	.44	.460	.538	.475
Rotated Blocks	.372	.396	.390	5.74	6.07	6.05	2.73	3.05	2.88	.59	.31	.20	.08	-.62	-.64	.669	.703	.678
General Science	.294	.282	.300	5.77	5.86	6.19	3.07	3.13	3.26	1.27	.85	.98	2.48	.64	1.94	.642	.651	.656
Hidden Figures	.500	.419	.464	7.45	6.40	6.97	3.11	2.62	3.33	.46	.66	.32	.04	.69	-.40	.741	.616	.766

APPENDIX B. COMMON ITEM ANALYSES RESULTS

Table B-1. Subtest and Composite Means and Standard Deviations Comprised of Common Items in Preliminary Equating Study

	P1		Q1		Q2	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Subtests						
Verbal Analogies	6.50	2.21	6.60	2.16	6.61	2.20
Arithmetic Reasoning	6.93	3.12	6.81	3.00	7.07	3.14
Reading Comprehension	6.15	2.68	6.29	2.62	6.52	2.75
Data Interpretation	3.96	1.95	4.05	1.96	4.23	2.05
Word Knowledge	6.53	2.71	6.64	2.67	6.57	2.66
Math Knowledge	5.54	2.40	5.59	2.40	5.62	2.43
Mechanical Comprehension	2.85	1.69	2.94	1.70	2.86	1.70
Electrical Maze	3.04	1.73	3.14	1.65	3.27	1.72
Scale Reading	12.19	3.88	12.40	3.91	12.47	3.97
Instrument Comprehension	4.53	2.32	4.71	2.32	4.63	2.33
Block Counting	4.17	1.06	4.19	1.03	4.14	1.05
Table Reading	11.05	3.17	11.38	3.13	11.44	3.13
Aviation Information	4.85	2.45	4.81	2.47	4.79	2.50
Rotated Blocks	3.26	1.54	3.38	1.51	3.31	1.51
General Science	3.45	1.81	3.35	1.79	3.38	1.86
Hidden Figures	4.61	1.78	4.48	1.72	4.60	1.77
Composites						
Pilot	49.20	13.30	50.20	12.95	50.26	13.10
Navigator Technical	61.09	17.57	61.75	17.10	62.44	17.72
Academic Aptitude	35.64	12.43	36.00	12.05	36.64	12.61
Verbal	19.19	6.61	19.54	6.48	19.71	6.66
Quantitative	16.44	6.70	16.45	6.48	16.92	6.85

Table B-2. Subtest and Composite Means and Standard Deviations Comprised of Common Items in Operational Equating Study

	P1		Q1		Q2	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Subtests						
Verbal Analogies	6.01	2.07	6.15	2.02	6.17	2.03
Arithmetic Reasoning	6.41	2.84	6.22	2.82	6.49	2.83
Reading Comprehension	5.92	2.34	5.98	2.41	6.29	2.35
Data Interpretation	3.86	1.79	3.91	1.86	4.13	1.83
Word Knowledge	6.28	2.65	6.26	2.58	6.25	2.55
Math Knowledge	5.22	2.26	5.23	2.18	5.30	2.16
Mechanical Comprehension	2.51	1.64	2.52	1.64	2.57	1.63
Electrical Maze	2.86	1.58	2.89	1.52	3.02	1.61
Scale Reading	11.21	3.61	11.51	3.63	11.65	3.56
Instrument Comprehension	4.01	2.18	4.05	2.22	4.01	2.20
Block Counting	4.06	1.17	4.09	1.10	4.09	1.12
Table Reading	10.98	2.86	11.15	2.75	11.30	2.74
Aviation Information	3.82	2.30	3.77	2.38	3.79	2.33
Rotated Blocks	2.97	1.50	3.05	1.50	3.07	1.49
General Science	3.04	1.70	2.97	1.72	3.04	1.72
Hidden Figures	4.31	1.65	4.14	1.62	4.27	1.62
Composites						
Pilot	45.48	11.35	46.17	11.31	46.63	11.12
Navigator Technical	57.47	15.00	57.75	14.86	58.98	14.72
Academic Aptitude	33.72	10.45	33.79	10.39	34.66	10.18
Verbal	18.21	5.96	18.41	5.90	18.73	5.78
Quantitative	15.50	5.77	15.38	5.78	15.93	5.73

APPENDIX C: PRELIMINARY CONVERSION TABLES FOR AFOQT FORMS Q1 AND Q2

Table C-1. Forms Q1 and Q2 Preliminary Conversion Table for Pilot Composite

Raw score	Percentile		Raw score	Percentile		Raw score	Percentile		Raw score	Percentile	
	Q1	Q2		Q1	Q2		Q1	Q2		Q1	Q2
0-41	1	1	83	15	16	124	48	51	165	94	92
42	2	1	84	16	17	125	50	52	166	94	94
43	2	1	85	17	17	126	51	53	167	95	94
44	2	2	86	17	18	127	52	54	168	96	95
45	2	2	87	17	19	128	53	55	169	96	95
46	2	2	88	18	20	129	54	56	170	97	96
47	2	2	89	19	20	130	55	57	171	97	96
48	2	2	90	20	20	131	56	58	172	97	96
49	3	2	91	20	20	132	57	60	173	97	97
50	3	3	92	20	21	133	58	61	174	98	97
51	3	3	93	21	22	134	60	62	175	98	97
52	3	3	94	22	23	135	61	63	176	98	97
53	3	3	95	23	24	136	62	63	177	98	98
54	3	3	96	24	24	137	63	64	178	99	98
55	3	3	97	24	25	138	63	65	179	99	98
56	4	4	98	25	26	139	64	66	180	99	98
57	4	4	99	26	27	140	65	67	181	99	98
58	4	4	100	27	28	141	66	69	182	99	99
59	4	4	101	27	28	142	67	70	183	99	99
60	5	5	102	28	29	143	69	71	184	99	99
61	5	5	103	28	30	144	70	73	185	99	99
62	6	6	104	29	31	145	73	74	186	99	99
63	6	6	105	30	32	146	74	75	187	99	99
64	6	6	106	31	33	147	75	76	188	99	99
65	6	6	107	32	34	148	76	77	189	99	99
66	7	7	108	33	35	149	77	78	190	99	99
67	7	7	109	34	36	150	78	79	191	99	99
68	7	7	110	35	37	151	79	80	192	99	99
69	8	8	111	36	38	152	80	81	193	99	99
70	8	8	112	37	39	153	81	82	194	99	99
71	8	8	113	38	41	154	83	83	195	99	99
72	9	9	114	39	42	155	84	84	196	99	99
73	10	10	115	41	42	156	84	84	197	99	99
74	10	10	116	41	43	157	85	85	198	99	99
75	11	11	117	42	43	158	86	86	199	99	99
76	11	11	118	42	44	159	86	86	200	99	99
77	12	12	119	43	45	160	88	87	201	99	99
78	12	13	120	44	46	161	89	88	202	99	99
79	13	13	121	45	47	162	90	89	203	99	99
80	13	13	122	46	48	163	91	90	204	99	99
81	13	14	123	47	50	164	93	91	205	99	99
82	14	15									

Table C-2. Forms Q1 and Q2 Preliminary Conversion Table for Navigator-Technical Composite

Raw score	Percentile		Raw score	Percentile		Raw score	Percentile		Raw score	Percentile		Raw score	Percentile	
	Q1	Q2		Q1	Q2		Q1	Q2		Q1	Q2		Q1	Q2
0-56	1	1	98	11	11	140	32	32	182	65	64	224	96	94
57	2	2	99	11	11	141	33	33	183	66	65	225	96	95
58	2	2	100	12	12	142	34	34	184	67	65	226	96	95
59	2	2	101	12	12	143	35	35	185	68	66	227	97	95
60	2	2	102	12	12	144	36	36	186	69	67	228	97	96
61	2	2	103	13	13	145	36	36	187	70	68	229	97	96
62	2	2	104	13	13	146	37	36	188	71	69	230	97	96
63	2	2	105	14	14	147	38	37	189	72	70	231	98	96
64	2	2	106	14	14	148	38	38	190	73	71	232	98	97
65	2	2	107	14	14	149	39	38	191	74	72	233	99	97
66	2	2	108	15	15	150	40	39	192	74	73	234	99	97
67	2	2	109	15	15	151	41	40	193	75	73	235	99	97
68	2	2	110	15	15	152	41	41	194	76	74	236	99	98
69	3	3	111	16	16	153	42	42	195	77	74	237	99	98
70	3	3	112	16	16	154	43	43	196	78	75	238	99	98
71	3	3	113	17	17	155	43	43	197	79	76	239	99	99
72	3	3	114	17	17	156	43	43	198	79	77	240	99	99
73	3	3	115	18	18	157	44	43	199	80	78	241	99	99
74	3	3	116	18	18	158	45	44	200	81	79	242	99	99
75	4	4	117	18	18	159	46	45	201	81	79	243	99	99
76	4	4	118	19	18	160	47	46	202	82	80	244	99	99
77	4	4	119	19	19	161	48	47	203	83	81	245	99	99
78	4	4	120	20	20	162	49	48	204	84	81	246	99	99
79	5	5	121	20	20	163	50	49	205	85	82	247	99	99
80	5	5	122	21	21	164	51	50	206	86	83	248	99	99
81	5	5	123	21	21	165	52	50	207	86	84	249	99	99
82	5	5	124	22	22	166	52	51	208	87	85	250	99	99
83	6	6	125	23	23	167	53	52	209	87	86	251	99	99
84	6	6	126	23	23	168	53	52	210	88	86	252	99	99
85	7	7	127	24	23	169	54	53	211	89	87	253	99	99
86	7	7	128	25	24	170	55	54	212	89	87	254	99	99
87	7	7	129	25	25	171	56	55	213	90	88	255	99	99
88	7	8	130	25	25	172	57	56	214	90	88	256	99	99
89	8	8	131	26	26	173	58	57	215	91	89	257	99	99
90	8	8	132	27	27	174	59	58	216	92	89	258	99	99
91	8	8	133	28	28	175	60	59	217	93	90	259	99	99
92	9	9	134	29	29	176	61	60	218	93	91	260	99	99
93	9	9	135	29	29	177	62	61	219	94	91	261	99	99
94	9	9	136	30	29	178	63	62	220	94	92	262	99	99
95	9	9	137	30	30	179	63	63	221	95	93	263	99	99
96	10	10	138	31	30	180	64	63	222	95	93	264	99	99
97	10	10	139	32	31	181	65	63	223	96	94	265	99	99

Table C-3. Forms Q1 and Q2 Preliminary Conversion Table for Academic Aptitude Composite

Raw score	Percentile Q1	Percentile Q2	Raw score	Percentile Q1	Percentile Q2	Raw score	Percentile Q1	Percentile Q2
0-24	1	1	67	16	16	109	63	61
25	1	2	68	16	16	110	65	62
26	1	2	69	17	16	111	67	63
27	1	2	70	18	17	112	68	65
28	1	2	71	18	18	113	69	67
29	2	2	72	19	18	114	70	68
30	2	2	73	20	18	115	71	69
31	2	2	74	21	19	116	72	70
32	2	2	75	21	20	117	76	71
33	2	3	76	22	21	118	78	72
34	2	3	77	23	21	119	79	75
35	2	3	78	24	22	120	80	76
36	3	3	79	25	23	121	81	78
37	3	3	80	26	24	122	82	79
38	3	3	81	27	25	123	83	81
39	3	4	82	28	26	124	84	82
40	3	4	83	28	27	125	85	83
41	4	4	84	29	28	126	86	84
42	4	5	85	31	28	127	87	85
43	5	5	86	33	29	128	88	86
44	5	5	87	34	31	129	89	87
45	5	5	88	36	33	130	90	88
46	5	5	89	37	34	131	91	89
47	5	6	90	38	35	132	92	90
48	6	6	91	38	36	133	93	91
49	6	6	92	40	37	134	93	92
50	6	7	93	41	38	135	94	93
51	7	7	94	43	38	136	95	93
52	7	7	95	44	40	137	95	95
53	8	8	96	45	41	138	96	95
54	9	9	97	47	43	139	96	96
55	9	9	98	49	44	140	97	96
56	9	9	99	50	45	141	97	97
57	9	9	100	51	47	142	98	97
58	10	10	101	52	49	143	98	98
59	10	10	102	53	50	144	99	98
60	11	10	103	54	51	145	99	99
61	11	11	104	54	52	146	99	99
62	12	11	105	57	53	147	99	99
63	13	12	106	59	54	148	99	99
64	14	13	107	61	57	149	99	99
65	15	14	108	62	59	150	99	99
66	16	15						

Table C-4. Forms Q1 and Q2 Preliminary Conversion Table for Verbal Composite

Raw score	Percentile Q1	Percentile Q2	Raw score	Percentile Q1	Percentile Q2
0-15	1	1	46	41	41
16	1	2	47	44	44
17	2	3	48	46	46
18	3	3	49	48	48
19	3	4	50	50	50
20	4	5	51	53	53
21	5	5	52	55	55
22	6	6	53	57	57
23	7	7	54	60	62
24	8	8	55	62	64
25	9	9	56	67	67
26	10	10	57	69	69
27	11	11	58	72	72
28	12	11	59	74	74
29	13	12	60	77	77
30	14	13	61	78	78
31	15	14	62	81	81
32	17	15	63	84	86
33	18	17	64	86	87
34	19	18	65	87	90
35	21	19	66	90	92
36	23	21	67	92	93
37	24	23	68	93	96
38	26	24	69	96	97
39	27	26	70	97	98
40	30	27	71	98	99
41	32	30	72	99	99
42	33	32	73	99	99
43	36	33	74	99	99
44	38	38	75	99	99
45	40	40			

Table C-5. Forms Q1 and Q2 Preliminary Conversion Table for Quantitative Composite

Raw score	Percentile Q1	Percentile Q2	Raw score	Percentile Q1	Percentile Q2
0-12	1	1	44	38	33
13	1	2	45	41	34
14	2	2	46	43	34
15	2	3	47	43	38
16	3	3	48	45	41
17	3	3	49	48	43
18	3	4	50	48	43
19	4	5	51	52	45
20	5	6	52	52	48
21	6	6	53	54	52
22	8	8	54	57	52
23	8	8	55	59	54
24	9	9	56	61	57
25	10	10	57	64	59
26	11	11	58	66	61
27	11	11	59	69	64
28	14	14	60	71	66
29	15	15	61	75	69
30	17	15	62	76	71
31	17	17	63	78	75
32	19	17	64	80	76
33	21	19	65	85	78
34	21	21	66	86	80
35	24	21	67	90	82
36	26	24	68	91	85
37	26	24	69	92	90
38	28	26	70	94	91
39	31	26	71	95	92
40	31	28	72	97	94
41	33	31	73	98	97
42	34	31	74	99	98
43	34	33	75	99	99

APPENDIX D: OPERATIONAL CONVERSION TABLES FOR AFOQT FORMS Q1 AND Q2

Table D-1. Forms Q1 and Q2 Operational Conversion Table for Pilot Composite

Raw score	Percentile		Raw score	Percentile		Raw score	Percentile		Raw score	Percentile	
	Q1	Q2		Q1	Q2		Q1	Q2		Q1	Q2
0-43	1	1	84	17	14	125	55	54	166	94	93
44	2	1	85	17	15	126	55	55	167	94	94
45	2	1	86	18	17	127	56	56	168	95	94
46	2	1	87	19	17	128	57	57	169	95	95
47	2	1	88	20	18	129	58	58	170	95	95
48	2	2	89	20	19	130	60	60	171	96	96
49	3	2	90	20	20	131	61	61	172	96	96
50	3	2	91	21	20	132	62	62	173	96	96
51	3	2	92	22	20	133	63	63	174	97	96
52	3	2	93	23	21	134	63	63	175	97	97
53	3	3	94	24	22	135	64	64	176	97	97
54	3	3	95	24	23	136	65	65	177	97	97
55	3	3	96	25	24	137	66	66	178	97	97
56	4	3	97	26	24	138	67	67	179	98	97
57	4	3	98	27	25	139	69	69	180	98	97
58	4	3	99	28	26	140	70	70	181	98	98
59	4	3	100	28	28	141	71	71	182	98	98
60	5	4	101	29	28	142	73	73	183	98	98
61	5	4	102	30	29	143	74	74	184	98	98
62	6	4	103	32	30	144	75	74	185	98	98
63	6	5	104	33	31	145	76	75	186	99	98
64	6	5	105	34	32	146	77	76	187	99	98
65	6	6	106	35	33	147	78	77	188	99	99
66	7	6	107	36	34	148	79	78	189	99	99
67	7	6	108	37	35	149	80	79	190	99	99
68	7	6	109	38	36	150	81	80	191	99	99
69	8	7	110	39	37	151	82	81	192	99	99
70	8	7	111	41	38	152	83	82	193	99	99
71	8	7	112	42	39	153	84	83	194	99	99
72	10	8	113	42	41	154	84	84	195	99	99
73	10	8	114	43	42	155	84	84	196	99	99
74	11	9	115	44	42	156	85	85	197	99	99
75	11	10	116	45	43	157	86	86	198	99	99
76	12	10	117	46	44	158	86	86	199	99	99
77	12	11	118	47	45	159	87	87	200	99	99
78	13	11	119	48	46	160	88	88	201	99	99
79	13	12	120	50	47	161	89	89	202	99	99
80	13	12	121	51	50	162	90	89	203	99	99
81	14	13	122	52	51	163	91	90	204	99	99
82	15	13	123	53	52	164	92	91	205	99	99
83	16	13	124	54	53	165	93	92			

Table D-2. Forms Q1 and Q2 Operational Conversion Table for Navigator-Technical Composite

Raw score	Percentile		Raw score	Percentile		Raw score	Percentile		Raw score	Percentile		Raw score	Percentile	
	Q1	Q2		Q1	Q2		Q1	Q2		Q1	Q2		Q1	Q2
0-62	1	1	103	12	11	144	38	36	185	73	69	226	96	95
63	2	1	104	12	11	145	38	36	186	73	70	227	96	95
64	2	1	105	13	12	146	39	37	187	74	71	228	96	96
65	2	1	106	13	12	147	40	38	188	74	72	229	96	96
66	2	2	107	14	12	148	41	38	189	75	73	230	96	96
67	2	2	108	14	13	149	42	39	190	76	73	231	97	96
68	2	2	109	15	13	150	43	40	191	77	74	232	97	96
69	2	2	110	15	14	151	43	41	192	78	74	233	97	97
70	2	2	111	15	14	152	43	42	193	79	75	234	97	97
71	2	2	112	16	15	153	44	43	194	79	76	235	97	97
72	2	2	113	16	15	154	45	43	195	80	77	236	97	97
73	2	2	114	17	15	155	46	43	196	81	78	237	98	97
74	3	2	115	17	16	156	47	44	197	81	79	238	98	98
75	3	2	116	18	16	157	49	45	198	82	79	239	98	98
76	3	2	117	18	17	158	50	45	199	83	80	240	98	98
77	3	3	118	18	17	159	51	46	200	83	81	241	99	98
78	3	3	119	19	18	160	52	47	201	84	81	242	99	99
79	4	3	120	20	18	161	52	48	202	85	82	243	99	99
80	4	3	121	21	18	162	53	49	203	86	83	244	99	99
81	4	3	122	21	19	163	54	50	204	86	83	245	99	99
82	4	3	123	22	20	164	55	51	205	87	84	246	99	99
83	5	4	124	23	20	165	56	52	206	87	85	247	99	99
84	5	4	125	23	21	166	57	52	207	88	86	248	99	99
85	5	4	126	24	21	167	58	53	208	88	86	249	99	99
86	5	5	127	25	22	168	59	54	209	88	87	250	99	99
87	6	5	128	25	23	169	60	55	210	89	87	251	99	99
88	6	5	129	26	23	170	61	56	211	89	88	252	99	99
89	7	5	130	27	24	171	62	57	212	90	88	253	99	99
90	7	6	131	28	25	172	63	58	213	90	89	254	99	99
91	7	6	132	29	26	173	63	59	214	91	89	255	99	99
92	8	7	133	29	27	174	64	60	215	91	90	256	99	99
93	8	7	134	30	28	175	65	61	216	92	90	257	99	99
94	8	7	135	30	29	176	65	62	217	93	91	258	99	99
95	9	8	136	31	29	177	65	63	218	93	91	259	99	99
96	9	8	137	32	30	178	66	63	219	94	92	260	99	99
97	9	8	138	33	30	179	67	64	220	94	93	261	99	99
98	9	9	139	34	31	180	68	65	221	94	93	262	99	99
99	10	9	140	35	32	181	69	65	222	95	94	263	99	99
100	11	9	141	36	33	182	70	66	223	95	94	264	99	99
101	11	10	142	36	34	183	71	67	224	95	94	265	99	99
102	12	10	143	37	35	184	72	68	225	95	95			

Table D-3. Forms Q1 and Q2 Operational Conversion Table for Academic Aptitude Composite

Raw score	Percentile Q1	Percentile Q2	Raw Score	Percentile Q1	Percentile Q2	Raw score	Percentile Q1	Percentile Q2
0-28	1	1	69	18	16	110	67	62
29	2	1	70	19	17	111	68	63
30	2	1	71	20	18	112	69	65
31	2	1	72	21	18	113	70	67
32	2	2	73	21	19	114	71	68
33	2	2	74	22	20	115	72	69
34	2	2	75	23	21	116	75	70
35	3	2	76	24	21	117	76	71
36	3	2	77	25	22	118	76	72
37	3	2	78	26	23	119	78	75
38	3	2	79	27	24	120	79	76
39	3	3	80	28	25	121	80	78
40	4	3	81	28	26	122	81	79
41	4	3	82	29	27	123	82	80
42	5	3	83	31	28	124	83	81
43	5	3	84	33	28	125	84	82
44	5	4	85	34	29	126	85	83
45	5	4	86	35	31	127	86	84
46	6	5	87	36	33	128	87	85
47	6	5	88	37	34	129	88	86
48	6	5	89	38	35	130	89	87
49	7	5	90	38	36	131	90	88
50	7	6	91	40	37	132	91	90
51	8	6	92	41	38	133	92	91
52	9	7	93	43	38	134	93	92
53	9	7	94	44	40	135	93	93
54	9	8	95	45	41	136	94	93
55	9	9	96	47	43	137	95	94
56	10	9	97	49	44	138	95	95
57	10	9	98	50	45	139	96	95
58	11	9	99	51	47	140	96	96
59	11	10	100	52	49	141	97	97
60	12	10	101	53	50	142	97	97
61	13	11	102	54	51	143	98	98
62	14	11	103	54	52	144	98	98
63	15	12	104	57	53	145	99	99
64	16	13	105	59	54	146	99	99
65	16	14	106	61	54	147	99	99
66	16	15	107	62	57	148	99	99
67	17	16	108	63	59	149	99	99
68	18	16	109	65	61	150	99	99

Table D-4. Forms Q1 and Q2 Operational Conversion Table for Verbal Composite

Raw score	Percentile Q1	Percentile Q2	Raw score	Percentile Q1	Percentile Q2
0-15	1	1	46	44	41
16	2	1	47	46	44
17	3	2	48	48	46
18	3	3	49	50	48
19	4	3	50	50	50
20	5	4	51	53	53
21	6	5	52	55	55
22	7	6	53	57	57
23	8	7	54	60	60
24	9	8	55	62	62
25	10	9	56	64	64
26	11	10	57	67	67
27	12	11	58	72	69
28	13	12	59	74	74
29	14	13	60	77	77
30	15	14	61	78	78
31	17	15	62	81	81
32	18	15	63	84	84
33	19	17	64	86	86
34	21	18	65	87	87
35	23	19	66	90	90
36	24	23	67	92	92
37	26	24	68	93	93
38	27	26	69	96	97
39	30	27	70	97	98
40	32	30	71	98	99
41	33	32	72	99	99
42	36	33	73	99	99
43	38	36	74	99	99
44	40	38	75	99	99
45	41	40			

Table D-5. Forms Q1 and Q2 Operational Conversion Table for Quantitative Composite

Raw score	Percentile Q1	Percentile Q2	Raw score	Percentile Q1	Percentile Q2
0-13	1	1	45	41	34
14	2	1	46	43	38
15	2	1	47	43	41
16	3	2	48	45	43
17	3	2	49	48	43
18	3	3	50	52	43
19	4	3	51	52	45
20	5	3	52	54	48
21	6	4	53	57	52
22	8	5	54	59	52
23	8	6	55	61	54
24	9	8	56	64	57
25	10	8	57	66	59
26	11	9	58	69	61
27	11	11	59	71	64
28	14	11	60	75	66
29	15	14	61	76	69
30	17	15	62	78	71
31	17	17	63	80	75
32	19	17	64	82	76
33	21	19	65	82	78
34	21	19	66	85	80
35	24	21	67	86	85
36	26	21	68	90	86
37	26	24	69	91	90
38	28	26	70	92	91
39	31	26	71	94	92
40	31	28	72	95	95
41	33	31	73	97	97
42	34	31	74	98	98
43	34	33	75	99	99
44	38	34			