AFHRL-TP-86-13

(12)

EQUIVALENCE OF COMPUTER AND PAPER-AND-PENCIL ARMED SERVICES VOCATIONAL APTITUDE BATTERY TESTS

Gerard L. Kiely
Anthony R. Zara
David J. Weiss

Department of Psychology
University of Minnesota
75 East River Road
Minneapolis, Minnesota 55455-0344

MANPOWER AND PERSONNEL DIVISION
Brooks Air Force Base, Texas 78235-5601

August 1986

Final Technical Paper for Period April 1984 - January 1986

AIR FORCE HUMAN RESOURCES LABORATORY

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235-5601

NOTICE

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 Public Affairs Office 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.

WILLIAM E. ALLEY, Scientific Advisor
Manpower and Personnel Division

RONALD L. KERCHNER, Colonel, USAF
Chief, Manpower and Personnel Division

8c. ADDRESS (City, State, and ZIP Code)
Brooks Air Force Base, Texas 78235-5601

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
62730F	7719	18	37

11. TITLE (Include Security Classification)
Equivalence of Computer and Paper-and-Pencil Armed Services Vocational Aptitude Battery Tests

12. PERSONAL AUTHOR(S)
Kiely, Gerard L.; Zara, Anthony R.; and Weiss, David J.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM Apr 84 TO Jan 86	August 1986	86

16. SUPPLEMENTARY NOTATION

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP
05	09	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
abilities testing
aptitude testing
Armed Services Vocational Aptitude Battery (ASVAB) (Continued)

19. ABSTRACT (Continue on reverse if necessary and identify by block number)
In order to investigate the possibility of obtaining equivalent results between computer and paper-and-pencil test administration, selected subtests of the Armed Services Vocational Aptitude Battery (ASVAB) were programmed for cathode-ray-tube (CRT) administration. Speeded subtest items were presented two ways: one at a time, and in blocks of several items. Paragraph comprehension items, which were too long to fit on the CRT screen, were presented in three different scrolling modes. The graphical portions of items were displayed from edited code created by a commercial digitizer. These CRT subtests were then compared with their paper-and-pencil counterparts using Air Force recruits in a counterbalanced design. Results indicated that obtaining equivalence between paper-and-pencil and computer administration appears feasible. Graphical items present the least difficulty. Item production for one type of speeded subtest was best approximated by single-item CRT presentation, and for another type of speeded subtest by CRT presentation of blocks of items. Additional research is required for the "too long" paragraph comprehension items in which more practice with the scrolling may be useful.

Item 18 (Concluded):

cathode ray tube (CRT)
classification
computer adaptive testing
computer testing
selection

AFHRL Technical Paper 86-13 August 1986

EQUIVALENCE OF COMPUTER AND PAPER-AND-PENCIL ARMED SERVICES VOCATIONAL APTITUDE BATTERY TESTS

Gerard L. Kiely
Anthony R. Zara
David J. Weiss

Department of Psychology
University of Minnesota
75 East River Road
Minneapolis, Minnesota 55455-0344

MANPOWER AND PERSONNEL DIVISION
Brooks Air Force Base, Texas 78235-5601

Reviewed by

Malcolm James Ree
Senior Scientist, Force Acquisition Branch

Submitted for publication by

Lonnie D. Valentine, Jr.
Chief, Force Acquisition Branch

SUMMARY

The armed services have a goal of administering their operational selection and classification enlisted test, the Armed Services Vocational Aptitude Battery (ASVAB), on the computer. Intentions are to administer the same kind of subtests as exist in the present operational paper-and-pencil (P&P) battery and to be able to do so in such a manner as to make it totally irrelevant whether an examinee receives the computer-administered subtest or the P&P version.

Several subtests consist of items which by their very nature would seem to be potential problems for transfer to the cathode-ray tube (CRT). There are two different types of speeded subtests (numerical operations and coding speed), whose items are easy and on which the final score generally depends on the number of items accomplished in a short time limit. The CRT is much too small to duplicate a full page of items such as those presented in a P&P mode. Similarly, the size of paragraph comprehension items prevents their appearing fully on the CRT screen. Finally, three of the ASVAB subtests have items with illustrations, and CRT presentation must switch from flat ink drawings to vertical light drawings on the screen.

Speeded subtests were programmed for two different CRT presentation modes, paragraph comprehension for three different modes, and graphics were displayed from code created by an off-the-shelf commercial digitizer. These CRT subtests were then compared with their P&P counterparts using Air Force recruits in a counterbalanced design. Results indicated that obtaining equivalence between P&P and computer administration appears feasible. Graphical items present the least difficulty. Item production for one type of speeded subtest was best approximated by single-item CRT presentation, and for another type of speeded test by CRT presentation of blocks of items. Additional research is required for the "too long" paragraph comprehension items in which more practice with the computer scope may be useful.

PREFACE

Several Armed Services Vocational Aptitude Battery (ASVAB) subtests consist of items whose characteristics were expected to change when administered on a computer screen. This technical paper examines several ways to administer these items in order to find one which is equivalent to paper-and-pencil administration. This effort is ancillary to the Air Force responsibility for item pools for a computer adapted ASVAB.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

I. INTRODUCTION

As the United States Armed Services move toward the implementation of computer adaptive testing (CAT) with the Armed Services Vocational Aptitude Battery (ASVAB), several major issues must be addressed to ensure the continuity and well-being of the military testing program. At the present time, the transition from traditional paper-and-pencil testing to computerized testing is being envisioned as a gradual process; therefore, the equivalence of the two coexisting methods of testing is of concern, because the resulting scores will be required to function interchangeably (Green, Bock, Humphreys, Linn, & Reckase, 1982).

The first question of equivalence that must be addressed is that of the influence of the presentation medium itself on otherwise identical tests. That is, if the ASVAB in its present form is administered by computer to a group of examinees, how would the resulting scores differ from those that would be obtained by testing in the paper-and-pencil medium? More specifically, three questions can be raised:

1. How would the scores differ by subtest?
2. What might account for some of these differences?
3. For some types of items, which of several modes of presentation on the computer screen might serve to minimize differences?

The present effort was designed to address these questions in two ways. One way that the equivalence of two procedures can be determined is by a comparison of the mean scores resulting from them to see if significant differences can be detected. Toward this end, the experimental design provided separate conditions for the study of alternative computer procedures against their paper-and-pencil counterparts. In addition, controls were built into the study to account for possible form and version differences, as well as practice effects. Methods of comparing mean differences were then applied.

The second approach was that of employing correlational methods. These techniques, which attempt to assess the degree of similarity rather than the extent of differences, included test-retest and internal consistency reliability comparisons across conditions, and exploratory and confirmatory factor analytic studies to compare factor structures.

For the purposes of this effort, 6 of the 10 ASVAB subtests were chosen for administration. The selection was based on some characteristic or set of characteristics of the subtests that might be expected to interact with the medium of presentation in either a positive or negative way, or that might require some change or modification from the previous paper-and-pencil form of presentation. These subtests can be divided into three groups on the basis of the criteria for selection: Paragraph Comprehension (PC), because of the problems involved in presenting it by computer that result from its unique item format; Numerical Operations (NO) and Coding Speed (CS), due to the speeded nature of these subtests; and Auto and Shop Information (AS), Mechanical Comprehension (MC), and Electronics Information (EI), because of their emphasis on graphical images, as well as items with no graphical content.

II. METHOD

Test Battery

This study used subtests of Forms 11, 12, and 13 of the ASVAB (Prestwood, Vale, Massey, & Welsh, 1985). The subtests used are shown in Table 1. The six subtests chosen for study were administered in two versions, A and B, for each of the three forms for a total of six alternatives for each subtest: 11A, 11B, 12A, 12B, 13A, and 13B. For the CS, AS, MC, and EI subtests, the two versions differ only in the ordering of the items; by contrast, the PC and NO subtests use different items in each version. All subtests consist of different items in each form. Scoring was based on number-correct raw scores for each subtest.

Table 1. ASVAB Subtests Used with Corresponding Number of Items, Operational Time in Minutes, and Speededness

Subtest	Number of items	Time allotted	Speeded
Paragraph Comprehension (PC)	15	13	No
Numerical Operations (NO)	50	3	Yes
Coding Speed (CS)	84	7	Yes
Auto and Shop Information (AS)	25	11	No
Mechanical Comprehension (MC)	25	19	No
Electronics Information (EI)	20	9	No

Experimental Design

Subjects and Data Collection

The initial sample consisted of 1,024 Air Force recruits distributed over 30 independent groups with repeated measures for each examinee. The 30 groups were subgroups of three general experimental groups: (a) Group 1, which was first administered the ASVAB tests by computer and then by paper-and-pencil; (b) Group 2, which took the tests first by paper-and-pencil and then by computer; and (c) Group 3 which took the ASVAB tests twice by paper-and-pencil.

After removal of 27 examinees with incomplete data, the final data consisted of item responses and subtest scores for 997 Air Force recruits on six subtests of the ASVAB, each taken twice in alternate forms, for a total of 12 subtest scores per examinee. Randomization was assured by preprinting 1,200 cards with coded conditions and ordering them in 40 groups, each group containing conditions 1 through 30 in sequence. Examinees were then assigned to groups by having them line up upon their arrival at the test center and handing out the cards in order until all examinees were accounted for. In this way, every 30th person was assigned to the same condition. All examinees were male recruits (to avoid possible confounding of results due to gender differences), and all were in their sixth day of basic military training at the time of testing. Testing of examinees occurred at the Air Force Human Resources Laboratory test facility at Lackland AFB, Texas, over a 2-week period during February 1985.

Once conditions were assigned, Group 3 (the group receiving paper-and-pencil tests in both sessions) was removed to a separate building for testing and was not exposed to the computerized test environment. Group 1 was brought into a room containing 30 carrels with Terak microcomputers, 28 of which were used for testing. Examinees were assigned randomly to computers. Due to computer memory limitations, each computer was programmed to administer only version A or version B tests. Within version, however, each machine contained all forms and conditions for all six subtests. The test each examinee received was determined by the condition code printed on his condition card, which was entered into the computer before the start of testing. This random assignment of examinees to conditions and to computers controlled for variations in screen resolution between computer monitors--a particularly important consideration for subtests containing graphical images.

Before testing commenced, Group 1 received a standard set of instructions, over earphones from trained test administrators, covering both the tests themselves and the operation of the computer; then, graphical demonstrations and user-paced exercises on computer keyboard use and the 10 keys required for testing were administered. All examinees also received standard ASVAB test instructions. Administrators were available throughout testing to answer questions and to help with problems. Upon completion of testing, examinees were allowed to leave quietly. During the first session, Group 2 was administered equivalent paper-and-pencil tests in another room. Following a break at the end of Session 1, Groups 1 and 2 switched places; i.e., Group 2 was administered instructions and testing by computer while Group 1 was administered paper-and-pencil tests.

Tests and Conditions

PC subtest. This subtest poses two problems when converted from paper-and-pencil to computer administration. First, some of the paragraphs used in the paper-and-pencil test are too long to fit on a cathode-ray-tube (CRT) screen at one time. Second, most of the reading comprehension paragraphs in the ASVAB tests are accompanied by multiple questions. Consequently, in the computer presentation medium, three different methods of presenting these kinds of items on a CRT (explained further on p. 4) were evaluated to determine which gave results most similar to those of paper-and-pencil administration.

NO and CS subtests. In the case of these highly-speeded, low-difficulty subtests, items are typically presented in groups in the paper-and-pencil medium, with instructions to answer as many items as possible within the time limit. A similar approach was taken with the computer presentation of these subtests by presenting several items on the screen at a time. This condition was compared with a second condition that presented items on the screen one at a time, to determine which item presentation condition was more similar to the paper-and-pencil administration.

AS, MC, and EI subtests. These subtests consist of both standard multiple-choice test items and multiple-choice items that use graphical images to describe physical, mechanical, or electronic concepts or components about which the examinee is questioned. They were presented on the computer screen in a single computer-presentation mode very similar to the presentation of their paper-and-pencil counterparts, to determine if differences resulted from

the translation of standard multiple-choice items and graphical images from the paper-and-pencil medium to a CRT.

Treatment Groups

The overall data collection plan, with the total numbers of examinees per condition (before eliminating those with incomplete data), is shown in Figures 1 through 3. Each figure is divided into sections representing the treatment groups to which the examinees were randomly assigned. A major advantage of this design was that it allowed the computer versus paper-and-pencil test administration variable to be examined both between and within subjects for all six subtests, thus allowing greater flexibility of analysis.

PC subtest. The data collection plan for the PC subtest is shown in Figure 1. To overcome the problem of lengthy PC items, a scrolling procedure was devised that enabled examinees to move forward and backward through the item text a line at a time. Scrolling was activated by depressing a designated forward or backward scrolling key which erased the text from the current screen (field) and replaced it with a different screen (field) of text. Depending on the scrolling key chosen (forward or backward), examinees were able to view new or previously presented material.

Three computer-administration Mode conditions (CRT-1 to CRT-3) were included to assess the effectiveness of three possible solutions to this complex item presentation problem. In the first Mode condition (CRT-1) the paragraph under consideration could be scrolled on the screen while one question at a time appeared beneath it in a separate, nonscrollable field. Previous questions about the paragraph were erased following an examinee's response and before the next question appeared, and examinees were unable to refer to previous paragraphs or previous questions pertaining to the current paragraph.

The second Mode condition (CRT-2) presented the paragraph and all questions related to it in a single scrollable field with no access to previous paragraphs or their questions. This condition allowed the examinee to use the entire screen to view the paragraph or single questions, and to move back and forth between the current paragraph and its questions. The third Mode condition (CRT-3) allowed the current paragraph to be scrolled on the screen, and beneath it questions could be scrolled separately. The displayed paragraph would change automatically as the items were scrolled to remain current with the displayed item, so that an examinee could scroll back to any previous paragraph and its questions. This condition also provided an answer-sheet type of display at the top of the screen which allowed examinees to monitor their progress through the test, and permitted them to go back and change answers to all previous PC items at any time throughout the test.

Each presentation condition appeared twice in the first two experimental groups, once with Version A and alternatively with Version B, for a total of six conditions with 50 to 60 examinees per condition. Due to the varied experimental conditions and limitations in sample size, it was necessary to limit the administration of the PC subtest to Form 11 of the ASVAB in Groups 1 and 2. Group 3, the paper-and-pencil only group, was not restricted by the three computer conditions; so the three forms of the PC subtest, Versions A and B, were administered instead, for a total of six conditions with 50 to 60 examinees per condition.

Group 1 (N=338): Computer-P&P

	Session 1		Session 2	
S_1	CRT1 11A	N=59	P&P 11B	N=59
.	CRT1 11B	N=56	P&P 11A	N=56
.	CRT2 11A	N=54	P&P 11B	N=54
.	CRT2 11B	N=56	P&P 11A	N=56
.	CRT3 11A	N=57	P&P 11B	N=57
S_{338}	CRT3 11B	N=56	P&P 11A	N=56

Group 2 (N=339): P&P-Computer

	Session 1		Session 2	
S_{339}	P&P 11A	N=56	CRT1 11B	N=56
.	P&P 11B	N=57	CRT1 11A	N=57
.	P&P 11A	N=58	CRT2 11B	N=58
.	P&P 11B	N=57	CRT2 11A	N=57
.	P&P 11A	N=56	CRT3 11B	N=56
S_{677}	P&P 11B	N=55	CRT3 11A	N=55

Group 3 (N=347): P&P-P&P

	Session 1		Session 2	
S_{678}	P&P 11A	N=58	P&P 11B	N=58
.	P&P 11B	N=58	P&P 11A	N=58
.	P&P 12A	N=58	P&P 12B	N=58
.	P&P 12B	N=58	P&P 12A	N=58
.	P&P 13A	N=58	P&P 13B	N=58
S_{1024}	P&P 13B	N=57	P&P 13A	N=57

Figure 1. Data Collection Design for PC Subtest.

NO and CS subtests. Figure 2 describes the data collection plan for the NO and CS subtests. Here, two computer presentation Mode conditions were included (CRT-1 and CRT-2) which, when combined with Versions A and B of each subtest, provided four conditions per group with 80 to 85 examinees per condition. In the CRT-1 condition, the items appeared on the CRT one at a time, and the examinees were instructed to answer as many of the items as possible within the allotted time. In the CRT-2 condition, several items were displayed at a time on the screen with the same instructions; for NO, three items appeared on the screen at the same time, and for CS, two blocks of seven items were displayed. Response times were recorded at both the single- and multiple-item level. Again, limitations in sample size restricted the administra-

Group 1 (N=338): Computer-P&P

	Session 1		Session 2	
S_1	CRT1 12A	N=83	P&P 12B	N=83
.	CRT1 12B	N=86	P&P 12A	N=86
.	CRT2 12A	N=87	P&P 12B	N=87
.	CRT2 12B	N=82	P&P 12A	N=82
S_{338}				

Group 2 (N=339): P&P-Computer

	Session 1		Session 2	
S_{339}	P&P 12A	N=85	CRT1 12B	N=85
.	P&P 12B	N=84	CRT1 12A	N=84
.	P&P 12A	N=85	CRT2 12B	N=85
.	P&P 12B	N=85	CRT2 12A	N=85
S_{677}				

Group 3 (N=300): P&P-P&P

	Session 1		Session 2	
S_{678}	P&P 11A	N=58	P&P 11B	N=58
.	P&P 11B	N=58	P&P 11A	N=58
.	P&P 12A	N=58	P&P 12B	N=58
.	P&P 12B	N=58	P&P 12A	N=58
.	P&P 13A	N=58	P&P 13B	N=58
.	P&P 13B	N=57	P&P 13A	N=57
S_{1024}				

Figure 2. Data Collection Design for NO and CS Subtests.

tion of the NO and CS subtests to Form 12 in Groups 1 and 2. Group 3 received both versions of all three forms for a total of six conditions with 50 to 60 examinees per condition.

AS, MC, and EI subtests. The data collection plan for the AS, MC, and EI subtests is shown in Figure 3. Unlike the other subtests, these consisted of only a single computer presentation condition that presented all items including graphical images. This permitted the administration of Forms 11, 12, and 13 to all three groups, for a total of six conditions per group with 50 to 60 examinees per condition. The graphical images for the computer-administered ASVAB tests were developed by the four-step process of (a) digitizing, (b) editing, (c) moving, and (d) concatenation and indexing. The digitizing process was done using a Kurta Series Two 12" x 17" digitizer pad.

The digitizer program that was written for this project (called DIGITIZER) allowed for digitizing to a half screen while scaling the image to make it as large as possible. In this way, the limited 120 vertical dots by 320 horizontal dots half-screen graphics window was used as completely as possible. DIGITIZER allows points, lines, and connected lines ("connect the dots" or polygon mode) to be plotted, either entered individually or "streamed" together.

After digitizing, all images were cleaned up and completed using the graphics editor. The graphics editor used was a reworked version of the Terak graphics editor GREDIT. It allows lines, points, circles, arcs, and arrow heads to be drawn or erased, or text to be entered or edited. It also allows the superimposing of one image upon another. Ultimately, all images were superimposed in this fashion, with one image on the top of the screen and the other on the bottom.

Once the image was refined to the desired level, it was centered within a half-screen area and located in the proper half screen (top or bottom), using a program called MOVER, which also allowed the duplication of one part of the screen on another part of the screen. The graphical portions of the screens were then concatenated and indexed into a random access file that was utilized by the test administration program to retrieve graphical images associated with text segments.

Data Analysis

Combined Conditions

The experimental design plan provided for three test administration groups: Group 1, to be tested by computer followed by paper-and-pencil; Group 2, tested by paper-and-pencil followed by computer; and Group 3, tested by paper-and-pencil twice. Within each group the three forms and two versions of each subtest were administered, except for PC in the first two groups where the three screen Modes (CRT-1 to CRT-3) were substituted for three forms and all examinees received Form 11. This resulted in 18 experimental conditions (3 groups x 3 forms x 2 versions). In addition, for the computer administration in Groups 1 and 2, two separate screen conditions were administered for the NO and CS subtests, thus adding an additional 12 conditions, yielding a total of 30 experimental conditions.

Group 1 (N=338): Computer-P&P

	Session 1		Session 2	
S_1	CRT 11A	N=59	P&P 11B	N=59
.	CRT 11B	N=56	P&P 11A	N=56
.	CRT 12A	N=54	P&P 12B	N=54
.	CRT 12B	N=56	P&P 12A	N=56
.	CRT 13A	N=57	P&P 13B	N=57
S_{338}	CRT 13B	N=56	P&P 13A	N=56

Group 2 (N=339): P&P-Computer

	Session 1		Session 2	
S_{339}	P&P 11A	N=56	CRT 11B	N=56
.	P&P 11B	N=57	CRT 11A	N=57
.	P&P 12A	N=58	CRT 12B	N=58
.	P&P 12B	N=57	CRT 12A	N=57
.	P&P 13A	N=56	CRT 13B	N=56
S_{677}	P&P 13B	N=55	CRT 13A	N=55

Group 3 (N=347): P&P-P&P

	Session 1		Session 2	
S_{678}	P&P 11A	N=58	P&P 11B	N=58
.	P&P 11B	N=58	P&P 11A	N=58
.	P&P 12A	N=58	P&P 12B	N=58
.	P&P 12B	N=58	P&P 12A	N=58
.	P&P 13A	N=58	P&P 13B	N=58
S_{1024}	P&P 13B	N=57	P&P 13A	N=57

Figure 3. Data Collection Design for AS, MC, and EI Subtests.

Analysis of Variance (ANOVA)

The model chosen for the ANOVA was a repeated measures multivariate (MANOVA) design using total subtest scores as dependent variables, with univariate follow-up ANOVAs for all significant effects. A multivariate analysis was selected because it offers a substantial advantage with respect to power, while controlling the Type I error rate over all subtests simultaneously at the nominal alpha level (Morrison, 1976).

A fundamental assumption underlying any MANOVA analysis is that of homogeneity of the covariance matrices within each cell. However, the MANOVA model is widely recognized to be robust to violation of this assumption, especially for large sample sizes. Furthermore, multivariate tests of homogeneity are very powerful, given large samples, often leading to a rejection of the hypothesis of equal covariance matrices on the basis of minor differences (Cooley & Lohnes, 1971). Univariate ANOVA models in a similar way assume homogeneity of variances, and in a like manner, are largely robust to violations of this assumption. This situation holds at least for between-subjects factorial models, but for repeated measures designs violations of the equal variance assumption have been shown to lead to a greatly increased probability of Type I error (O'Brien & Kaiser, 1985). Unfortunately, there is no known test at present for the homogeneity of repeated measures variances. Therefore, the best defense against an increased incidence of Type I error is to interpret significant outcomes accordingly, when the variances over sessions appear to be heterogeneous.

MANOVA is based upon a comparison of the latent structures of the between-groups sums of squares and cross-products (SSCP) matrix, $\underline{H}$, and the within-groups SSCP matrix of subtest scores, $\underline{E}$, for a given effect. This is accomplished through the decomposition of the product matrix, $\underline{H}\underline{E}^{-1}$, and the analysis of the resulting latent roots. Although there is no uniformly most powerful statistic for conducting this analysis, all of the widely accepted statistics are some function of the latent roots of $\underline{H}\underline{E}^{-1}$, and the choice of an appropriate statistic is dependent upon the rank of the matrix, or the number of significant roots obtained.

For the situation where the rank of the matrix is 1, the most powerful test of significance has been shown to be Roy's Largest Root test; whereas, when the rank is greater than 1, the Pillai-Bartlett Trace V test is most powerful (Olson, 1976). The Pillai-Bartlett offers the additional advantage of being the most robust to the violation of the assumptions of the MANOVA model, whereas Roy's test is extremely sensitive to violation of the homogeneity assumption.

For the present analyses, four multivariate tests of significance were applied: Roy's Largest Root, Hotelling's T, Wilks' Lambda, and the Pillai-Bartlett Trace V test. For every experimental outcome, the results of all four tests were in complete agreement. Therefore, the Pillai-Bartlett Trace V statistic, which has been recommended for general use (Olson, 1976) is reported for the MANOVA analyses.

The Pillai-Bartlett Trace V test is

$$V = \sum_{i}^{s} C_i/(1+C_i) \qquad (1)$$

where C_i is the ith latent root of the matrix $\underline{H}\underline{E}^{-1}$ and S is its rank. Pillai derived an approximation to the F distribution (Morrison, 1976) for V as

$$F = \frac{(dfe - p + S)V}{b(S - V)} \qquad (2)$$

with Sb and S(dfe - p + s) degrees of freedom, where dfe is the degrees of freedom within groups, p is the number of dependent variables, S is the rank of $\underline{H}\underline{E}^{-1}$, and b is the larger of p and the degrees of freedom between groups.

The follow-up analysis for those MANOVA outcomes found to be significant was based on univariate ANOVA F tests to identify the specific subtests for which the effects were significant. Once a MANOVA result is found significant, however, all follow-up statistical tests are unconstrained in terms of Type I error. That is, the Type I error rate applies at the individual test level, and over numerous tests, the rate is compounded. Caution must be exercised in such a case in interpreting significant outcomes. Furthermore, standard forms of post hoc statistical tests, while varying in power and flexibility, are all increasingly sensitive with increasing sample sizes, and for the present study, relatively large samples were required to adequately provide for the correlational analyses, such as factor analysis. Such large samples serve to increase the sensitivity to post hoc statistical tests to the point where meaningless differences become prominent and obscure the important group differences under experimental manipulation.

Paper-and-pencil baseline analysis. The purpose of the first analysis was to examine the effects across all six subtests of (a) Form--the three test forms--ASVAB Forms 11, 12, and 13; (b) Version--the two Versions (A and B) within each form; and (b) Session--the repeated measurement of paper-and-pencil tests. This analysis was confined to Group 3 (N = 333), which functioned as a baseline comparison group, being tested entirely in the paper-and-pencil medium. Figure 4 shows the experimental design plan for Group 3.

The analysis consisted of a repeated measures MANOVA with univariate ANOVA follow-up tests for all significant effects across session within subjects, and test form and test version between subjects. The dependent variables were the total scores for each of the six subtests taken by each examinee (PC, NO, CS, AS, MC, and EI).

Paragraph Comprehension. The second analysis was designed to determine if differences were observed in mean test scores among the three modes of computer presentation of the PC subtest between subjects, and between the computer and paper-and-pencil presentation media both between and within subjects. Test form differences were also examined to eliminate confounding with the other effects.

Figure 5 shows the experimental design plan for the PC analysis. All examinees were tested using Form 11 of the PC subtest to allow sufficient sample size for a fully crossed design. Group 1, consisting of 332 examinees, was subdivided into three groups of approximately equal size. Each subgroup was administered one test version in one mode of the computer medium in the first testing session, followed by paper-and-pencil testing with the alternate version in the second session; subgroups of between 50 and 60 examinees within each mode subgroup were administered Versions A and B of ASVAB Form 11 PC in

Test version presentation order	Examinees	Session 1 Medium	Session 1 Form & version	Session 1 Subtest	Session 2 Medium	Session 2 Form & version	Session 2 Subtest
1	665 to 720	P&P	11A N=56	AS, MC, EI	P&P	11B N=56	PC, NO, CS, AS, MC, EI
	721 to 776		12A N=56	PC, NO, CS, AS, MC, EI		12B N=56	PC, NO, CS, AS, MC, EI
	777 to 833		13A N=57	PC, NO, CS, AS, MC, EI		13B N=57	PC, NO, CS, AS, MC, EI
2	834 to 889	P&P	11B N=56	PC, NO, CS, AS, MC, EI	P&P	11A N=56	PC, NO, CS, AS, MC, EI
	890 to 941		12B N=52	PC, NO, CS, AS, MC, EI		12A N=52	PC, NO, CS, AS, MC, EI
	942 to 997		13B N=56	PC, NO, CS, AS, MC, EI		13B N=56	PC, NO, CS, AS, MC, EI

Figure 4. Experimental Design Plan for Paper-and-Pencil Baseline Analysis.

counterbalanced order. Group 2 (N = 332) was tested with the presentation media in reverse order, i.e., paper-and-pencil followed by computer presentation.

The method of analysis was a repeated measures MANOVA with the PC subtest score as the sole dependent variable. The effects that were examined as between-subjects variables for the PC analysis included three modes of computer presentation (Mode), computer versus paper-and-pencil presentation (Medium), and test version (Version). The effect examined as a within-subjects variable was the computer versus paper-and-pencil presentation medium across sessions (Session).

Speeded tests. The purpose of the third analysis was to determine if differences existed between mean scores resulting from the two modes of computer presentation of the NO and CS subtests, and between the computer and paper-and-pencil presentation media both between and within subjects. Test version differences were examined to determine if they interacted with these primary effects.

Figure 6 summarizes the experimental design plan for these analyses. All

Medium presentation order	Examinees	Session 1 Mode	Session 1 Form & version	Session 1 Subtest	Session 2 Mode	Session 2 Form & version	Session 2 Subtest
1	1 to 58	CRT-1	11A N=58	PC	P&P	11B N=58	PC
	59 to 114		11B N=56	PC		11A N=56	PC
	115 to 168	CRT-2	11A N=54	PC	P&P	11B N=54	PC
	169 to 223		11B N=55	PC		11A N=55	PC
	224 to 278	CRT-3	11A N=55	PC	P&P	11B N=55	PC
	279 to 332		11B N=54	PC		11A N=54	PC
2	333 to 387	P&P	11A N=55	PC	CRT-1	11B N=55	PC
	388 to 441		11B N=54	PC		11A N=54	PC
	442 to 499	P&P	11A N=58	PC	CRT-2	11B N=58	PC
	500 to 553		11B N=54	PC		11A N=54	PC
	554 to 609	P&P	11A N=56	PC	CRT-3	11B N=56	PC
	610 to 664		11B N=55	PC		11A N=55	PC

Figure 5. Experimental Design Plan for Analysis of Paragraph Comprehension Subtest.

Medium presentation order	N	Session 1 Mode	Session 1 Form & version	Session 1 Subtest	Session 2 Mode	Session 2 Form & version	Session 2 Subtest
1	1 to 83	CRT-1	12A N=83	NO, CS	P&P	12B N=83	NO, CS
	84 to 168		12B N=85	NO, CS		12A N=85	NO, CS
	169 to 252	CRT-2	12A N=84	NO, CS	P&P	12B N=84	NO, CS
	253 to 332		12B N=80	NO, CS		12A N=80	NO, CS
2	333 to 417	P&P	12A N=85	NO, CS	CRT-1	12B N=85	NO, CS
	418 to 500		12B N=83	NO, CS		12A N=83	NO, CS
	501 to 584	P&P	12A N=84	NO, CS	CRT-2	12B N=84	NO, CS
	585 to 664		12B N=80	NO, CS		12A N=80	NO, CS

Figure 6. Experimental Design Plan for Analysis of Speeded Subtests.

testing was limited to Form 12 to allow sufficient sample size for a fully crossed design. Group 1, with 332 examinees, was subdivided randomly into two subgroups of approximately equal size. Each subgroup was administered one test in one of the two computer Modes (single-item or multiple-item screen) in the first testing session, followed by paper-and-pencil testing with the alternate version of Form 12 in the second session. Group 2, also with 332 examinees, was tested in the reverse order, with paper-and-pencil testing in the first session and computer testing in the second session. Within each of the Mode subgroups of Group 1 and Group 2, half of the examinees were assigned to either Version A or Version B of ASVAB Form 12 in random order to test the Version effect. This yielded a total of eight experimental groups of 80 to 85 examinees each. (Due to sample size limitations, within-subject test version and medium effects were necessarily combined into a single experimental condition.)

The method of analysis was a repeated measures MANOVA with the NO and CS subtest scores as the dependent variables, with univariate ANOVA follow-up tests for all significant tests. The between-subjects variables were test version (Version), computer presentation mode (Mode), and computer versus paper-and-pencil presentation medium (Medium), and the within-subjects variable was the computer versus paper-and-pencil medium (Session) effect.

Graphical subtests. The fourth analysis was designed to examine the differences between the computer and paper-and-pencil presentation media for the AS, MC, and EI subtests. These differences were analyzed for Forms 11, 12, and 13 of each subtest and by version. Medium and Version effects were examined in a completely unconfounded design. For the graphical subtests, a particular question of interest is what differences in performance, if any, result from the presentation of graphical images on a computer screen versus the printed page. The analysis of the MC subtest addresses this question directly, since of the 25 items comprising the test, 23 for Form 11, 21 for Form 12, and 22 for Form 13 contain graphical images. In addition, the MC subtests, as with all others, were randomly distributed over 28 computers to control for individual screen resolution differences.

Figure 7 shows the experimental design plan for these tests. The 332 examinees in Group 1 were randomly assigned to three subgroups. Each subgroup was administered one form of each subtest, one version in the computer medium in Session 1 and the alternate version in the paper-and-pencil medium in Session 2. The second group of 332 examinees was tested in a similar manner but with the order of presentation medium reversed.

The analysis was a repeated measures MANOVA with total subtest scores as dependent variables. The effects that were examined as between-subjects variables included test form, computer versus paper-and-pencil presentation medium, and test version. The effect examined as a within-subjects variable was the computer versus paper-and-pencil presentation medium.

Reliability Analyses

Experimental group and session abbreviations are presented in Figure 8. The first character stands for medium of administration: C for computer, P for paper-and-pencil administration; the second character denotes the experimental group (1, 2, or 3); the third digit indicates testing session 1 or 2. Due to the nature of the data collected, two types of reliability comparisons were made.

Test-retest. The first comparison used a test-retest correlation design and was computed for all the ASVAB subtests. The subtests CS, AS, MC, and EI contain the same items in both Version A and Version B, providing true test-retest data. However, subtests PC and NO use different items in Versions A and B, thus changing the comparison slightly to one of an alternate-forms correlation, although the versions are nominally parallel. The comparison of interest was that between Groups 1 and 2 (C11 then P12, and P21 then C22) versus the paper-and-pencil only (P31 then P32) group. In these comparisons the Group 3 reliabilities served as a baseline against which the other experimental conditions could be judged. The reliabilities were compared by t-tests for each contrast after performing Fisher's (1921) r to z transformation on

Medium presentation order	N	Session 1 Medium	Session 1 Form & version	Session 1 Subtest	Session 2 Medium	Session 2 Form & version	Session 2 Subtest
1	1 to 58	CRT	11A N=58	AS, MC, EI	P&P	11B N=58	AS, MC, EI
	59 to 114		11B N=56	AS, MC, EI		11A N=56	AS, MC, EI
	115 to 168		12A N=54	AS, MC, EI		12B N=54	AS, MC, EI
	169 to 223		12B N=55	AS, MC, EI		12A N=55	AS, MC, EI
	224 to 278		13A N=55	AS, MC, EI		13B N=55	AS, MC, EI
	279 to 332		13B N=54	AS, MC, EI		13A N=54	AS, MC, EI
2	333 to 387	P&P	N=55	AS, MC, EI	CRT	N=55	AS, MC, EI
	388 to 441		11B N=54	AS, MC, EI		11A N=54	AS, MC, EI
	442 to 499		12A N=58	AS, MC, EI		12B N=58	AS, MC, EI
	500 to 553		12B N=54	AS, MC, EI		12A N=54	AS, MC, EI
	554 to 609		13A N=56	AS, MC, EI		13B N=56	AS, MC, EI
	610 to 664		13B N=55	AS, MC, EI		13A N=55	AS, MC, EI

Figure 7. Experimental Design Plan for Analysis of Graphical Subtests.

		Test Administration	
		Session 1	Session 2
Experimental Group	1	C11	P12
	2	P21	C22
	3	P31	P32

Figure 8. Overview of Test Conditions by Group and Session for Computer (CRT) and Paper-and-Pencil (P&P) Groups.

the test-retest correlations.

Internal consistency. The final reliability comparisons were made using an internal consistency measure of reliability, Cronbach's (1951) coefficient alpha, which reduced to the Kuder-Richardson (1937) Formula 20 (KR-20) due to the dichotomous item responses. Because internal consistency reliability estimates are not appropriate for speeded tests, only subtests PC, AS, MC, and EI were considered in these comparisons.

The data collection plan provided the means to perform both matched groups (within-group across sessions) and multiple independent groups (across-groups within session) KR-20 comparisons. In order to perform the complete matched group comparison across all subtests by Form and Version for each experimental group, 72 tests of the type proposed by Feldt (1980) would be necessary. A simple independent groups comparison for each subtest by form and version within each session would require that 48 tests using Hakstian and Whalen's (1976) method be made. The sheer number of statistical tests necessary to adequately compare the KR-20 reliabilities suggested that the findings obtained would be extremely difficult to interpret and confounded by Type 1 error.

As an alternative, the KR-20 estimates were averaged separately for each subtest within each group and session (C11, P12, P21, C22, P31, and P32) and examined for differences. Collapsing the data in this way was justified because the ASVAB forms and versions are designed to yield similar measurement characteristics; thus, any important reliability differences across presentation medium or session would be consistent across all versions and forms.

Studying the reliabilities in this manner allowed the discovery of possible trends or differences in KR-20s both within groups across sessions, and across groups within sessions. As in the test-retest analysis, the paper-and-pencil-only groups (Group 3, P31 and P32) were used as a baseline against which the mean KR-20s from the other groups were judged.

Analysis of Structure

Between-subtest structure. In order to examine the possible effect of different item presentation methods on a battery of tests, an analysis was performed on the similarity of factor and covariance structures between subtests at the subtest score level, under both administration medium conditions. The data were correlation and covariance matrices for between-subtest scores computed from examinees' total scores for each of the six subtests.

One method used to determine the equivalence of test presentation media was the examination of the factor structure of the between-subtest correlations across conditions. For this comparison, unrotated principal factor analysis was used as the factoring method. Although sets of common factor loadings are usually not unique and many different sets of loading values can define a solution, principal factor analysis yields a factor solution that defines both a unique common factor space and a unique set of factor loadings (Harman, 1976). Therefore, factor loadings from the unrotated principal factor analysis solution (defined by extraction of maximum variance from the correlation matrices with squared multiple correlations on the diagonals) of the subtests were directly compared for equivalence across media of presentation.

The equality of between-subtest covariance matrices and factor structures was examined using the maximum likelihood methods of LISREL VI (Jöreskog & Sörbom, 1984). There are three main indices of general model fit yielded by LISREL. The first is the overall chi-square (χ^2). The test made by chi-square judges the fit of the constrained model against the alternative hypothesis that the estimated covariance matrix is unconstrained. Degrees of freedom for χ^2 are calculated by

$$df = \tfrac{1}{2}k(k + 1) - t \tag{3}$$

where k is the number of observed variables, and t is the number of free parameters estimated. Jöreskog and Sörbom (1984) suggest that χ^2 be used as an index of the degree of model fit and not strictly as a test statistic.

The second index of overall model fit is the goodness-of-fit index (GFI). GFI is defined by

$$GFI = 1 - \frac{tr(\hat{\Sigma}^{-1}S - I)^2}{tr(\hat{\Sigma}^{-1}S)^2} \tag{4}$$

where tr is the trace of the indicated matrices, $\hat{\Sigma}$ is the fitted matrix, S is the observed covariance matrix, and I is an identity matrix. The range of GFI is between zero and 1.0, and it is a measure of the amount of covariance and variance accounted for by the model.

The last model fit index is the root mean square residual (RMR) defined as

$$RMR = \left[2\sum_{i=1}^{k}\sum_{j=1}^{i}(s_{ij} - \hat{\sigma}_{ij})^2/k(k + 1)\right]^{1/2} \tag{5}$$

where k is the number of observed variables in the model, s_{ij} is the observed variance or covariance, and $\hat{\sigma}_{ij}$ is the estimated covariance or variance component. The interpretation of RMR depends on the relative sizes of the covariances and variances of the observed variables. For example, a large value for RMR as compared to the average observed covariance or variance would be an indication of a model that did not fit the data very well.

Jöreskog (1971) outlined a method by which the factor structures of test batteries could be tested for equality across different groups of examinees. Models assuming different levels of equalities were tested sequentially until the level of structured equality appropriate for the between-subtest data was found. The first model tested, the most strict test of covariance equality, assumed that, within each Session (1 and 2), the between-subtest scores covariance matrices for the paper-and-pencil administrations were equal for each group (Session 1: $\Sigma_{G2} = \Sigma_{G3}$ and Session 2: $\Sigma_{G1} = \Sigma_{G3}$). This test of the model that the covariance structures within the paper-and-pencil administrations were equal across groups provided a baseline against which the covariance structures of computer-administered tests could be judged. Next, the same test was made, again by session, with the addition of the covariances for the computer administrations ($\Sigma_{G1} = \Sigma_{G2} = \Sigma_{G3}$). These tests are generalizations of the Bartlett test for homogeneity of variances (Morrison, 1976) and are susceptible to having a high degree of power when sample sizes are large, causing rejection of the null hypothesis when minor differences are present (Cooley & Lohnes, 1971).

The next model tested, holding a much less strong equality, was that within each Session (1 and 2) the between-subtest scores yielded the same factor structure in each group.

Within-subtest structure. The analysis of the similarity of factor structures within subtests was also performed for subtests AS, MC, and EI. These were the only non-speeded subtests containing the same items in both versions (A and B). By combining examinees tested on either version, the sample size requirements of factor analysis were nearly met, with approximately 100 persons per subtest available across Forms (11, 12, and 13) within each cell (C11, P12, P21, C22, P31, P32) of the experimental design.

The factor structures were compared through uniterated principal factor analysis, as described previously, of the item intercorrelation matrices composed of phi correlations. Five unrotated principal factors were extracted from each correlation matrix. These comparisons were made across both subjects and media using the first session baseline paper-and-pencil group for all comparisons (C11 and C22 versus P31). The subtest factor structures from both computer-administered conditions were compared against the subtest factor structures of the Session 1 paper-and-pencil-only group (Group 3) because it yielded an adequate representation of the factor structures of the subtests in all the paper-and-pencil groups.

Item Analysis

Due to limitations in sample size (between 50 and 75 per cell), only classical test theory item parameters (point-biserial, biserial, and proportion correct) were calculated and analyzed for the non-speeded subtests. The sample size demands of item response theory item parameterization using LOGIST

(Wood, Wingersky, & Lord, 1976)--500 to 1,000 per cell (Wood & Lord, 1976)--greatly exceed the number of examinees in this study.

The subtests under study were chosen because they are the most problematic ASVAB subtests for computer administration. Of particular concern was the possible difference in measurement properties of items containing graphical images (in subtests AS, MC, and EI) due to problems in translating the image from paper to a computer CRT screen. This question of measurement equivalence was neatly addressed at the subtest level because of the distribution of graphical content items among the three tests. The MC subtest contains between 21 to 23 graphical items (23 in Form 11, 21 in Form 12, 22 in Form 13) out of the 25 total, whereas subtest EI has only two or three out of 20 items. Therefore, any computer administration effect on graphical items would impact on subtest MC in its entirety and cause mean differences across media, which would be found through the MANOVA analysis. Subtest EI, having almost no graphical content items, was used to index possible graphical item administration differences since it has similar test objectives, being a non-speeded, technical information test. Subtest AS was not used in this analysis because it contains several graphical items (5 in Form 13, and 6 in Forms 11 and 12), disqualifying it as either a high or low graphical content subtest.

Administration medium differences for graphical content items were also compared at the factor structure level. The differences in the unrotated principal factor solutions for subtest MC across media of presentation were compared with the factor structure differences for subtest EI across the same conditions. Any large discrepancy between MC and EI factor structure differences would provide evidence that computer administration changes the interrelationships among graphical items, implying that the translation of images from paper to CRT screen differentially affects items with graphical content.

III. RESULTS

Analysis of Variance

Paper-and-Pencil Baseline Analysis

Table 2 shows the outcome of the MANOVA for the paper-and-pencil baseline group, both within and across sessions, for all subtests. Significance ($p \leq .05$) is indicated for one within-session factor, test form (Form), the repeated measures factor (Session), and for the interaction of test version (Version) with Form and Session. All other factors and interactions were not significant; therefore, no further analyses were necessary.

The results of the univariate follow-up analyses by subtest for the significant effects identified in the multivariate baseline analysis are given in Table 3. For the Form factor, significance was found for the AS subtest. Significant Session effects are shown for the NO, CS, and EI subtests, and PC shows a significant outcome for the Version by Form by Session interaction.

The next step in the analysis was to examine the difference in the means for each level of each effect. Tables 4 through 9 show the means for all subtests by condition. A significant difference in the mean scores by Form was found for the AS subtest; the largest of these was between Form 11 and Form

Table 2. Pillai-Bartlett Trace Values, Degrees of Freedom, Approximate F Ratios, and Estimated Significance Levels (p) for the Baseline MANOVA (N = 333)

Source of variation	Trace V	Degrees of freedom Between	Within	F	p<
Between Groups					
Form	.10	12	646	2.96	.001**
Version	.01	6	322	.30	.935
Form x Version	.04	12	646	1.11	.353
Within Groups					
Session	.35	6	322	29.21	.001**
Form x Session	.05	12	646	1.45	.138
Version x Session	.02	6	322	.93	.473
Version x Form x Session	.08	12	646	2.30	.007**

**Statistically significant at p < .01.

Table 3. Results of Univariate ANOVAs for Significant Effects Identified in the Baseline MANOVA (N = 333)

Effect, degrees of freedom, and subtest	Mean squares Between	Within	F	p<
Form (2,327)				
PC	6.82	8.96	.76	.468
NO	160.97	97.17	1.66	.192
CS	317.30	303.33	1.05	.353
AS	129.51	39.02	3.32	.037*
MC	10.74	33.05	.33	.723
EI	20.22	23.13	.87	.418
Session (1,327)				
PC	.29	3.01	.10	.755
NO	312.22	14.92	20.93	.001**
CS	5506.34	34.10	161.47	.001**
AS	.38	1.18	.33	.568
MC	5.23	2.07	2.53	.113
EI	12.99	1.85	7.02	.008**
Version x Form x Session (2,327)				
PC	32.99	3.01	10.94	.001**
NO	26.34	14.92	1.77	.173
CS	14.35	34.10	.42	.657
AS	.37	1.18	.32	.728
MC	1.50	2.07	.73	.484
EI	1.13	1.85	.61	.542

*Statistically significant at p < .05.
**Statistically significant at p < .01.

Table 4. Mean and Standard Deviation of Scores, Kuder-Richardson Formula 20 Reliability Coefficient (KR20), and Number of Examinees (N), for the PC Subtest by Group and Session

Session 1						Session 2					
Mode	Version	Mean	SD	KR20	N	Mode	Version	Mean	SD	KR20	N
Group 1: Computer						**Group 1: Paper-and-Pencil**					
CRT-1	11A	10.93	2.26	.556	59	-	11B	11.38	2.57	.661	58
CRT-1	11B	9.77	2.41	.550	56	-	11A	11.93	2.90	.750	57
CRT-2	11A	8.07	3.31	.761	53	-	11B	11.31	2.26	.525	58
CRT-2	11B	8.84	3.05	.730	56	-	11A	11.86	2.75	.727	57
CRT-3	11A	10.07	2.33	.586	57	-	11B	11.93	2.48	.649	56
CRT-3	11B	9.95	2.38	.540	56	-	11A	11.82	2.45	.636	57
Group 2: Paper-and-Pencil						**Group 2: Computer**					
-	11A	12.14	2.66	.723	57	CRT-1	11B	12.11	2.44	.670	56
-	11B	12.25	2.34	.652	56	CRT-1	11A	12.26	2.18	.586	57
-	11A	11.71	2.72	.717	58	CRT-2	11B	10.71	3.55	.849	55
-	11B	11.78	2.19	.576	55	CRT-2	11A	9.82	4.00	.885	54
-	11A	12.55	2.37	.701	58	CRT-3	11B	12.68	2.25	.661	56
-	11B	12.55	1.95	.579	58	CRT-3	11A	12.69	1.81	.467	55
Group 3: Paper-and-Pencil						**Group 3: Paper-and-Pencil**					
-	11A	11.79	2.57	.702	58	-	11B	11.54	2.81	.716	57
-	11B	11.59	2.87	.752	58	-	11A	11.66	2.55	.672	58
-	12A	11.54	2.43	.637	56	-	12B	12.86	1.97	.595	58
-	12B	12.27	2.20	.627	56	-	12A	11.41	2.46	.633	54
-	13A	11.93	2.76	.750	55	-	13B	11.50	2.59	.653	58
-	13B	11.91	2.55	.675	57	-	13A	12.51	1.79	.470	57

Table 5. Mean and Standard Deviation of Scores, and Number of Examinees (N), for the NO Subtest by Group and Session

Session 1					Session 2				
Mode	Version	Mean	SD	N	Mode	Version	Mean	SD	N
Group 1: Computer					**Group 1: Paper-and-Pencil**				
CRT-1	12A	33.59	7.86	83	-	12B	38.71	7.44	87
CRT-2	12B	34.36	8.75	86	-	12A	38.16	7.79	87
CRT-1	12A	27.60	10.77	87	-	12B	37.75	7.35	85
CRT-2	12B	29.16	11.48	82	-	12A	37.93	7.78	84
Group 2: Paper-and-Pencil					**Group 2: Computer**				
-	12A	37.18	8.30	87	CRT-1	12B	36.96	7.15	85
-	12B	38.64	7.35	86	CRT-2	12A	36.92	6.89	84
-	12A	36.49	7.96	86	CRT-1	12B	31.88	9.29	85
-	12B	38.80	7.18	83	CRT-2	12A	31.04	11.62	85
Group 3: Paper-and-Pencil					**Group 3: Paper-and-Pencil**				
-	11A	40.21	7.85	58	-	11B	41.40	8.64	57
-	11B	40.14	7.41	58	-	11A	40.50	7.18	58
-	12A	38.75	7.91	56	-	12B	40.21	7.57	58
-	12B	38.50	8.05	56	-	12A	40.52	7.85	54
-	13A	40.36	7.47	55	-	13B	41.69	7.44	58
-	13B	40.11	7.16	57	-	13A	42.81	6.42	57

Table 6. Mean and Standard Deviation of Scores, and Number of Examinees (N), for the CS Subtest by Group and Session

Session 1					Session 2				
Mode	Version	Mean	SD	N	Mode	Version	Mean	SD	N
Group 1: Computer					**Group 1: Paper-and-Pencil**				
CRT-1	12A	70.78	11.44	83	-	12B	55.57	11.74	87
CRT-2	12B	72.83	10.12	86	-	12A	56.15	11.12	87
CRT-1	12A	49.05	12.14	87	-	12B	50.34	12.66	85
CRT-2	12B	51.30	13.27	82	-	12A	52.73	11.38	84
Group 2: Paper-and-Pencil					**Group 2: Computer**				
-	12A	52.02	11.24	87	CRT-1	12B	73.94	11.40	85
-	12B	53.03	10.94	86	CRT-2	12A	74.07	10.48	84
-	12A	53.07	13.28	86	CRT-1	12B	54.84	13.36	85
-	12B	52.39	12.23	82	CRT-2	12A	56.19	13.90	85
Group 3: Paper-and-Pencil					**Group 3: Paper-and-Pencil**				
-	11A	52.48	14.16	57	-	11B	58.07	14.43	57
-	11B	51.19	14.03	58	-	11A	56.14	15.18	58
-	12A	54.75	12.59	56	-	12B	60.95	13.22	58
-	12B	53.27	12.99	56	-	12A	59.61	12.95	54
-	13A	51.76	12.20	55	-	13B	58.84	10.92	58
-	13B	52.77	12.36	57	-	13A	58.09	14.22	57

Table 7. Mean and Standard Deviation of Scores, Kuder-Richardson Formula 20 Reliability Coefficient (KR20), and Number of Examinees (N), for the AS Subtest by Group and Session

Group and version	Session 1				Session 2			
	Mean	SD	KR20	N	Mean	SD	KR20	N
Group 1	**Computer**				**Paper-and-Pencil**			
11A/11B	17.02	4.64	.784	59	17.57	4.69	.798	58
11B/11A	17.63	4.87	.819	56	17.72	4.99	.823	57
12A/12B	17.35	4.56	.795	54	18.07	4.53	.800	58
12B/12A	16.75	5.32	.844	56	17.84	4.88	.830	57
13A/13B	17.12	4.94	.827	57	17.39	4.97	.832	55
13B/13A	17.04	5.17	.835	56	17.75	5.49	.861	57
Group 2	**Paper-and-Pencil**				**Computer**			
11A/11B	17.96	4.56	.795	57	18.16	4.62	.804	56
11B/11A	17.95	5.45	.865	55	18.70	4.57	.805	57
12A/12B	16.86	4.77	.804	58	17.50	4.45	.785	58
12B/12A	17.55	4.87	.822	55	17.44	4.85	.817	57
13A/13B	17.28	4.51	.781	58	17.30	4.77	.806	56
13B/13A	18.69	4.26	.784	58	18.51	4.27	.782	55
Group 3	**Paper-and-Pencil**				**Paper-and-Pencil**			
11A/11B	18.45	4.67	.823	57	18.60	4.03	.753	57
11B/11A	19.09	4.36	.797	58	19.16	4.26	.791	58
12A/12B	17.21	4.90	.825	56	17.52	4.80	.818	58
12B/12A	17.04	4.69	.806	56	17.48	4.36	.784	54
13A/13B	18.44	4.31	.775	55	18.16	4.57	.797	58
13B/13A	17.60	4.72	.807	57	17.74	4.96	.824	57

Table 8. Mean and Standard Deviation of Scores, Kuder-Richardson Formula 20 Reliability Coefficient (KR20), and Number of Examinees (N), for the MC Subtest by Group and Session

Group and version	Session 1				Session 2			
	Mean	SD	KR20	N	Mean	SD	KR20	N
Group 1	**Computer**				**Paper-and-Pencil**			
11A/11B	15.61	4.87	.792	59	16.67	4.39	.753	58
11B/11A	16.04	5.14	.822	56	17.05	4.62	.787	57
12A/12B	15.83	4.64	.777	54	17.09	4.90	.817	57
12B/12A	16.45	4.30	.746	56	17.98	4.25	.761	57
13A/13B	16.74	4.10	.739	57	17.27	4.30	.772	55
13B/13A	16.23	3.85	.675	56	17.33	3.78	.697	57
Group 2	**Paper-and-Pencil**				**Computer**			
11A/11B	18.04	3.76	.702	57	18.14	3.80	.706	56
11B/11A	18.16	3.70	.705	56	18.21	3.90	.725	57
12A/12B	16.84	4.93	.814	58	17.21	4.22	.748	58
12B/12A	16.98	4.14	.727	55	16.37	4.72	.787	57
13A/13B	17.31	3.65	.677	58	17.54	3.37	.627	56
13B/13A	17.50	3.56	.659	58	17.73	3.62	.673	55
Group 3	**Paper-and-Pencil**				**Paper-and-Pencil**			
11A/11B	17.48	4.40	.776	58	17.37	4.20	.755	57
11B/11A	17.72	3.96	.722	58	17.91	4.03	.737	58
12A/12B	16.98	4.61	.781	56	17.64	3.87	.696	58
12B/12A	16.88	4.80	.796	56	17.81	4.42	.779	54
13A/13B	17.75	3.60	.662	55	17.67	4.00	.726	58
13B/13A	16.88	4.32	.754	57	16.96	4.26	.745	57

Table 9. Mean and Standard Deviation of Scores, Kuder-Richardson Formula 20 Reliability Coefficient (KR20), and Number of Examinees (N), for the EI Subtest by Group and Session

Group and version	Session 1				Session 2			
	Mean	SD	KR20	N	Mean	SD	KR20	N
Group 1	**Computer**				**Paper-and-Pencil**			
11A/11B	12.71	3.56	.698	59	13.26	3.49	.695	58
11B/11A	13.13	3.83	.743	55	13.60	3.40	.687	57
12A/12B	13.09	4.16	.772	54	13.67	4.15	.785	57
12B/12A	13.59	3.94	.748	56	14.18	3.47	.688	57
13A/13B	13.40	3.71	.737	57	13.77	3.70	.754	55
13B/13A	13.46	3.32	.672	56	14.21	3.41	.709	57
Group 2	**Paper-and-Pencil**				**Computer**			
11A/11B	13.47	3.81	.751	57	14.07	3.81	.761	56
11B/11A	13.61	3.78	.755	56	13.51	3.71	.740	57
12A/12B	13.76	2.93	.560	58	13.81	2.99	.576	58
12B/12A	14.02	3.52	.705	55	13.44	4.23	.793	57
13A/13B	14.00	3.34	.686	58	14.05	3.32	.679	56
13B/13A	14.64	3.35	.709	58	14.45	3.08	.647	55
Group 3	**Paper-and-Pencil**				**Paper-and-Pencil**			
11A/11B	13.12	4.05	.769	58	13.68	3.51	.702	57
11B/11A	14.02	3.39	.701	58	14.19	3.64	.742	58
12A/12B	13.68	4.54	.818	56	14.22	4.05	.783	58
12B/12A	14.34	3.15	.639	56	14.59	3.11	.646	54
13A/13B	14.00	2.62	.504	55	14.34	2.75	.561	58
13B/13A	12.70	3.96	.751	57	13.30	3.66	.719	57

12, with a difference of 1.41, suggesting that no practical or meaningful significance can be attached to this outcome. Therefore, a conclusion of no important Form effect was drawn for the paper-and-pencil baseline analysis.

For the Session effect, the NO subtest showed an increase in mean score of 1.37 from Session 1 to Session 2, CS had a gain of 5.75, and EI differed by .28. Therefore, a very slight NO increase and a somewhat larger CS increase were implied by these statistically significant mean differences.

The PC subtest analysis indicated a significant Version by Form by Session interaction. However, for this subtest, the range in mean scores from the lowest to the highest over all 12 testing conditions was only 1.30. Because of the low magnitude of this mean difference, this interaction was interpreted as a psychometrically nonmeaningful difference.

Computer versus Paper-and-Pencil

PC subtest. The results of the MANOVA for the PC subtest for the computer and paper-and-pencil groups are reported in Table 10 (means and standard deviations by condition are in Table 4). The statistically significant between-subjects effects include the mode of computer administration (Mode) and computer versus paper-and-pencil administration (Medium). No significant between-subjects interactions were revealed. The within-subjects repeated measures factor (Session) was found to be significant, as well as the Medium by Session, Version by Mode by Session, Version by Medium by Session, and Mode by Medium by Session interactions.

The follow-up analysis for the PC subtest began with the Mode effect. For the three computer screen Modes, the mean scores were 10.32 for CRT-1, 8.46 for CRT-2, and 9.99 for CRT-3; the corresponding mean differences were 1.86 raw-score points between Modes CRT-1 and CRT-2, .33 between 1 and 3, and 1.53 between 2 and 3. This indicates that screen condition 2 was different from 1 and 3, whereas conditions 1 and 3 were not significantly different from each other.

For the computer versus paper-and-pencil Medium effect, the mean difference within Session between the paper-and-pencil group (mean = 12.19) and the computer group (mean = 9.59) was 2.60 points. For the Session effect, the first and second session mean difference was .82 points. Upon examination of the Medium by Session interaction, it was found that the group that took the computer test in session 2 scored 2.15 points higher (mean = 11.74) than the group that took it during the first session (mean = 9.59). For the paper-and-pencil tests, however, the second session mean score (mean = 11.67) was only .52 lower than the first session scores (mean = 12.19), a finding of little practical importance.

The Medium, Session, and Medium by Session effects were reevaluated with the low-scoring second presentation Mode group removed. This reduced the Medium mean difference to 2.03 points, and increased the Session difference for the computer groups slightly to 2.43 points.

Of the three-way interactions of Version by Mode by Session, Version by Medium by Session, and Mode by Medium by Session, all revealed only minor mean differences of less than 1 point across conditions, but the Mode by Session

Table 10. Results of the F-tests of Significance from the MANOVA for the Computer Versus Paper-and-Pencil Administration PC Subtest Conditions (N = 664)

Source of variation	Degrees of freedom	Mean square	F	p<
Between Subjects				
Version	1	.01	.00	.979
Mode	2	218.66	21.31	.001**
Medium	1	590.51	57.56	.001**
Version x Mode	2	.66	.06	.937
Version x Medium	1	.57	.06	.814
Mode x Medium	2	15.77	1.54	.216
Version x Mode x Medium	2	21.69	2.11	.122
Within Cells	652	10.26		
Within Subjects				
Session	1	221.47	65.88	.001**
Version x Session	1	.20	.06	.810
Mode x Session	2	2.47	.74	.480
Medium x Session	1	531.81	158.21	.001**
Version x Mode x Session	2	15.41	4.58	.011*
Version x Medium x Session	1	14.03	4.17	.041*
Mode x Medium x Session	2	90.48	26.92	.001**
Version x Mode x Medium x Session	2	4.88	1.4	.235
Within Cells	652	3.36		

*Statistically significant at $p < .05$.
**Statistically significant at $p < .01$.

interaction was a meaningful one in that Mode differences existed only within one of the two Medium conditions.

Speeded subtests. The results of the MANOVA for the speeded NO and CS subtests are shown in Table 11 (mean scores by condition are in Table 5 for NO and Table 6 for CS). For the between-subjects effects, Medium, Mode, and the Medium by Mode interaction were statistically significant. Within subjects, Session, Version by Session, Medium by Session, and Medium by Mode by Session effects were found to be statistically significant.

Table 12 contains the results of the analyses of statistically significant effects by subtest. Medium differences were indicated for the NO subtest, while Mode differences were observed for both NO and CS. The between-subjects Medium by Mode interaction was found significant for the CS subtest. Within subjects, the Session factor was significant for both NO and CS, and all of the interactions tested were significant for both subtests except for Version by Session, which was significant for NO but decidedly nonsignificant for CS.

Examination of the mean scores for the Medium effects shows that the paper-and-pencil group (mean NO = 37.79) obtained higher scores than the computer group (mean NO = 31.12) in Session 1 by 6.67 points for the NO subtest, and 8.22 (mean CS scores of 60.96 and 52.74, respectively) for CS; both differ-

Table 11. Pillai-Bartlett Trace V Values with 2 and 655 Degrees of Freedom, Approximate F Ratios, and Estimated Significance Levels (p) from the MANOVA for the Computer Versus Paper-and-Pencil Administration Conditions for the NO and CS Subtests (N = 664)

Source of variation	Trace V	F	p<
Between Groups			
Version	.00	1.11	.331
Medium	.01	3.29	.038*
Mode	.23	96.09	.001**
Version x Medium	.00	.82	.439
Version x Mode	.00	.11	.897
Medium x Mode	.01	3.30	.038*
Version x Medium x Mode	.00	.04	.965
Within Groups			
Session	.06	21.59	.001**
Version x Session	.01	3.54	.030*
Medium x Session	.63	556.01	.001**
Mode x Session	.00	.08	.925
Version x Medium x Session	.00	.84	.434
Version x Mode x Session	.00	1.43	.241
Medium x Mode x Session	.42	237.07	.001**
Version x Medium x Mode x Session	.00	.73	.482

*Statistically significant at $p < .05$.
**Statistically significant at $p < .01$.

ences were clearly consequential, even though the ANOVA analysis found the CS difference marginally statistically nonsignificant. In Session 2, the paper-and-pencil group (mean = 38.11) outscored the computer group (mean = 34.32) by 3.79 for NO, and by 11.29 for CS (means = 53.52 and 64.81). The CS result is clearly substantive. The Session effect overall, with a 1.77 point gain for NO (means = 34.46 and 36.23), and 2.32 points for CS (means = 56.91 and 59.23), does not indicate much of a real difference, since practice effects would be expected on these speeded items.

Upon comparison of computer screen Mode differences, it was found first that the CRT-1 Mode condition (mean = 34.02) produced higher scores than the CRT-2 condition (mean = 28.21) in Session 1 by 5.81 for NO, an important mean difference; and by 21.86 for CS (means = 71.89 and 50.03), a highly important difference. These disparities hardly diminished for the computer group in Session 2, with an NO difference of 5.42 (means = 37.03 and 31.61) and a CS difference of 18.45 (means = 74.04 and 55.59).

Further analyses of the computer versus paper-and-pencil medium differences by computer presentation mode within Session 1 revealed that, for NO, the CRT-1 condition (mean = 34.02) showed no practical difference from the paper-and-pencil condition (mean = 37.79), but that the CRT-2 condition (mean = 28.21) resulted in substantially lower scores than did both the paper-and-pencil condition (mean = 37.79) and the alternate computer Mode (mean = 34.02). The same results were obtained in Session 2, with the CRT-1 condition

Table 12. Results of Univariate ANOVAs for Significant Effects Identified in the MANOVA for the Subtests with 1 and 656 Degrees of Freedom (N = 664)

Effect and subtest	Mean squares Between	Within	F	p<
Medium				
NO	685.89	106.29	6.45	.011*
CS	782.94	228.41	3.43	.065
Mode				
NO	3,006.37	106.29	28.23	.001**
CS	40,997.32	228.41	179.49	.001**
Medium x Mode				
NO	14.35	106.29	.13	.713
CS	1,145.51	228.41	5.02	.025*
Session				
NO	1,028.29	40.18	25.59	.001**
CS	1,784.08	57.15	31.22	.001**
Version x Session				
NO	263.96	40.18	6.57	.011*
CS	.88	57.15	.02	.901
Medium x Session				
NO	9,083.82	40.18	226.09	.001**
CS	31,570.67	57.15	552.45	.001**
Medium x Mode x Session				
NO	2,256.84	40.18	56.17	.001**
CS	27,114.36	57.15	474.47	.001**

*Statistically significant at $p < .05$.
**Statistically significant at $p < .01$.

(mean = 37.03) not being different from the paper-and-pencil condition (mean = 38.11), but with both of these conditions revealing mean scores much higher than those of the CRT-2 condition (mean = 31.60). For CS, an outcome similar to that for the NO subtest was obtained, except that for Session 1 the CRT-1 computer presentation condition (mean = 71.88) had mean scores much higher than the paper-and-pencil (mean = 52.74) and the CRT-2 computer presentation (mean = 50.04) conditions. For Session 2, the CRT-1 condition (mean = 74.04) was again much higher than either the paper-and-pencil (mean = 53.52) or the CRT-2 (mean = 55.58) conditions.

Although a significant Version by Session interaction was reported by the ANOVA analysis for the NO subtest, pairwise comparisons of mean differences found no statistical significance. The largest difference discovered was that between Version A in Session 1 (mean = 33.64) and Version A in Session 2 (mean = 36.30), a difference of only 2.66 points. The Version B means were 35.26 for Session 1, increasing to 36.13 in Session 2. For the Medium by Session interaction, the difference between paper-and-pencil and computer administration scores decreased in Session 2 for NO from 6.67 to 3.79, and increased for CS from 8.22 to 11.28, neither by a significant amount.

Graphical subtests. Table 13 provides the results of the MANOVA analysis for the AS, MC, and EI graphical subtests (mean scores by condition for these tests are in Tables 5, 6, 7). This analysis revealed no statistically signif-

Table 13. Pillai-Bartlett Trace V Values, Degrees of Freedom, Approximate F Ratios, and Estimated Significance Levels (p) from the MANOVA for the Computer Versus the Paper-and-Pencil Administration Conditions for the AS, MC, and EI Subtests (N = 664)

Source of variation	Trace V	Degrees of freedom Between	Within	F	p<
Between Groups					
Version	.00	3	650	.32	.810
Form	.01	6	1,302	1.44	.196
Medium	.01	3	650	2.34	.072
Version x Form	.00	6	1,302	.36	.902
Version x Medium	.00	3	650	1.03	.379
Form x Medium	.01	6	1,302	1.37	.225
Version x Form x Medium	.00	6	1,302	.48	.827
Within Groups					
Session	.09	3	650	20.24	.001**
Version x Session	.00	3	650	.84	.474
Form x Session	.01	6	1,302	.82	.552
Medium x Session	.04	3	650	9.10	.001**
Version x Form x Session	.00	6	1,302	.35	.908
Version x Medium x Session	.01	3	650	1.88	.131
Form x Medium x Session	.01	6	1,302	1.45	.191
Version x Form x Medium x Session	.01	6	1,302	1.37	.226

**Statistically significant at p < .01.

Table 14. Results of Univariate ANOVAs for Significant Effects Identified in the MANOVA for the AS, MC, and EI Subtests with 1 and 652 Degrees of Freedom (N = 664)

Effect and subtest	Mean squares Between	Within	F	p<
Session				
AS	58.70	2.19	26.75	.001**
MC	122.61	2.93	41.88	.001**
EI	17.45	2.13	8.21	.004**
Medium x Session				
AS	3.89	2.19	1.77	.184
MC	70.74	2.93	24.17	.001**
EI	21.86	2.13	10.28	.001**

**Statistically significant at p < .01.

icant effects for the between-subjects factors, Version, Form, or Medium. Significance was indicated for the repeated measures Session effect, and the Medium by Session interaction. Table 14 shows the results of the univariate follow-up analyses for the statistically significant effects, revealing significance for all three graphical subtests on the Session effect, and for MC and EI on the Medium by Session interaction.

Comparison of the means by condition for the significant effects revealed small differences in means across session of .42 for AS, .61 for MC, and .23 for EI, certainly not of any psychometric consequence. For the Medium by Session effect for MC and EI, differences of a similar magnitude were revealed, with the MC computer scores increasing by 1.43 across sessions (means = 16.14 and 17.57), while the paper-and-pencil scores declined by .21 (means = 17.57 and 17.21). For the EI subtest, computer scores gained .66 across sessions (means = 13.23 and 13.89) while the paper-and-pencil mean score was .20 lower (means = 13.92 and 13.72), all relatively small changes.

Comparison of the means for the MC Session 1 Medium effect, specifically computer (mean = 16.14) versus paper-and-pencil (mean = 17.43) presentation of graphical image items, showed a mean difference of 1.29. Further inspection showed that the largest part of this difference was attributable to Form 11, where the paper-and-pencil group (mean = 18.07) outscored the computer group (mean = 15.82) by 2.25. For Form 12, the means were 16.17 for the computer group and 16.88 for the paper-and-pencil group, for a difference of .71; for Form 13, the means were 16.42 and 17.32, respectively, for a difference of .90.

Summary

No important psychometrically meaningful significant differences were demonstrated for the paper-and-pencil baseline analysis for any subtests, except for a practice effect on CS. For the computer versus paper-and-pencil equivalence analysis, the PC subtest revealed a major difference in the Mode factor, with the CRT-2 condition resulting in important mean differences from the CRT-1 and CRT-3 conditions. Even with the effects of this condition removed, however, differences were still shown between the computer and paper-and-pencil presentation media. It was also found that the group that took the computer test second scored higher on it than the group that took it first. However, the group that took the paper-and-pencil test second, did not obtain higher scores than the other group's first session paper-and-pencil test.

The speeded test comparisons revealed that for NO, the CRT-2 presentation mode was decidedly inferior in performance, while the CRT-1 condition did not differ significantly from the paper-and-pencil results. For CS, it was found that the CRT-1 computer condition resulted in higher scores than did either the CRT-2 or the paper-and-pencil conditions, which did not differ from each other.

The analysis for the graphical subtests revealed no psychometrically meaningful differences for any effects for any subtests, including the first session Medium effect for the MC test which addresses directly the question of differences due to paper-and-pencil versus computer presentation of graphical image test items.

Reliability Analysis

Test-Retest Correlations

Table 15 shows test-retest correlations (for CS, AS, MC and EI) or alternate forms retest correlations (for PC and NO) for each subtest. For the speeded subtests (NO and CS), all the correlations for Groups 1 and 2 for both computer administration modes were significantly lower (at the .05 level) than those from Group 3, except for the NO subtest in Group 2 computer-administered Mode 1. These significant differences in retest correlations were fairly large, the smallest being .20 for CS (.65 versus .85) in Group 2 Mode 2, increasing to .36 for NO (.33 versus .69) in Group 1 Mode 2.

Table 15. Test-Retest Correlations (r), for Subtests CS, AS, MC, and EI and Alternate Forms Retest Correlations for PC and NO Subtests and Number of Examinees (N) by Group and Mode

Subtest & mode/ version	Group 1		Group 2		Group 3	
	r	N	r	N	r	N
PC						
1	.42	114	.75*	109	.49	112
2	.45	109	.49	112		
3	.33	109	.79*	111		
NO						
1	.44*	168	.65	168	.69	108
2	.33*	164	.50*	164		
CS						
1	.57*	168	.56*	168	.85	108
2	.60*	164	.65*	164		
AS						
11	.91	114	.88*	109	.93	112
12	.93	109	.88*	112	.94	108
13	.89*	109	.95	111	.95	113
MC	.81*	332	.87	332	.88	333
EI	.79*	332	.89	332	.85	333

Note. Group 3 PC correlation is based on examinees taking Form 11 only, and the NO and CS correlations are based on examinees taking Form 12 only, in correspondence with the forms administered to Groups 1 and 2.

*Indicates that the Group 1 or Group 2 correlation is significantly different from the Group 3 correlation at the .05 level.

The graphical subtests (AS, MC, and EI) showed a similar pattern of test-retest correlations. For each graphical subtest, the Group 1 correlations were significantly lower than those from Group 3, although for AS only Form 13 yielded significantly lower correlations. The only Group 2 test-retest correlations significantly lower than Group 3 were from AS, Forms 11 and 12. Although these reliabilities (test-retest) were statistically significantly lower for the groups in which one test administration was computerized, the actual differences in test-retest correlations were not very large, ranging from .05 (AS in Group 2, Form 11) to .07 (MC in Group 1).

The pattern of test-retest correlations was somewhat different for subtest PC as there were no differences in test-retest correlations between Groups 1 and 3 that reached significance at the .05 level. The Group 2 PC test-retest correlations were significantly higher than Group 3 for Modes 1 and 3. These were the only test-retest correlations for any of the subtests that were significantly higher in either of the computer administration groups (1 and 2) than in the paper-and-pencil-only Group 3 (.75 and .79 versus .49).

Internal Consistencies

Mean KR-20s for each non-speeded subtest by Group and Session are in Table 16 (KR-20s for each group and condition are in Tables 4 through 9). Tables 4 through 9 show that the differences in KR-20s within cells were of the same magnitude as any difference found between cells; thus, the data were studied as means instead of as individual values from separate test administrations.

Table 16. Average Kuder-Richardson Formula 20 Reliability (KR20) for Subtests PC, AS, MC, and EI by Experimental Group and Session

	Session 1				Session 2			
Group	PC	AS	MC	EI	PC	AS	MC	EI
	Computer				Paper-and-Pencil			
Group 1	.558	.817	.758	.728	.674	.824	.764	.720
	Paper-and-Pencil				Computer			
Group 2	.658	.808	.714	.694	.596	.800	.711	.699
	Paper-and-Pencil				Paper-and-Pencil			
Group 3	.727	.806	.748	.697	.694	.794	.740	.692

Note. For the PC subtest in the computer conditions, only Modes 1 and 3 were included, due to the extreme Mode 2 differences. In the paper-and-pencil conditions, only Form 11 was included so as to correspond with the computer conditions.

The lowest mean reliabilities for both computer administration (.558 and .658) and paper-and-pencil (.727) were obtained for the PC subtest, as expect-

ed, due to its short length. Also, the largest differences in mean KR-20s across media of administration for any subtest were found for the reliabilities of PC, with differences of .100 (C11 versus P21) and .169 (C11 versus P31). The PC subtest also showed the only large increase in mean KR-20 across sessions, with the Group 1 mean KR-20 increasing from .558 in Session 1 to .674 in Session 2.

Subtest AS yielded much more consistent mean KR-20 values, with a range from .794 (Group 3, Session 2) to .824 (Group 1, Session 2). This .03 range is striking in that both values are from Session 2 paper-and-pencil-administered tests, implying that the computer versus paper-and-pencil comparisons were more equal, with only .009 (C11 versus P21) and .011 (C11 versus P31) differences in Session 1 and .024 (P12 versus C22) and .006 (C22 versus P32) differences in Session 2.

The Session 1 and 2 mean KR-20 differences were all smaller than .01 for subtest MC. This corresponds with the pattern found for subtest AS, where the largest difference was a .012 decrease from Session 1 to Session 2. Within Session 1, there was a larger mean KR-20 difference for MC between the two paper-and-pencil administrations (P21 versus P31, .03) than between C11 and P31 (.01), though all the differences were very small.

The mean KR-20 values for subtest EI conformed to the general pattern found for the other graphical subtests, indicating no effect on internal consistency for computer versus paper-and-pencil test administration. As found for MC, there was no computer administration effect within examinees for EI, with all differences in mean KR-20s across sessions being .008 or less. In both sessions, the mean KR-20 was higher in Group 1 than in the other groups, with the overall range being .692 (P32) to .728 (C11).

Analysis of Structure

Across Subtests

For the calculation of the between-subtest correlation and covariance matrices from the computer administrations, only scores obtained from examinees taking Modes 1 and 3 for the PC subtest and Mode 1 of the NO and CS subtests were included. The MANOVA analysis showed that the Mode 2 scores for PC, NO, and CS were much different than both Mode 1 and Mode 3 in terms of subtest means and variances, thus their exclusion from this analysis. Table 17 shows the correlations between subtest scores by Group and Session which were analyzed by principal factor analysis to yield the eigenvalues shown in Table 18. The eigenstructure is fairly similar within each cell (Group and Session combination), with a large first factor capturing between 63.4% (P21) and 69.8% (P32) of the common variance. All cells seem to conform to a two-factor solution except C11, for which three factors are indicated. Cell C11 is the only one where the third eigenvalue is positive, with the third factor accounting for 9.9% of the common variance. The third factor is present at the expense of Factor 2 which accounts for 13.3% less common variance than the Factor 2 in any other cell in Session 1.

Table 19 gives the factor loadings by Group and Session from the principal factor analyses. Factor 1 loads highest on, and is defined by, the graphical subtests (AS, MC, EI) in every cell except P31, where PC loads higher than AS (.572 versus .542). There are also substantial loadings for PC on

Table 17. Intercorrelations of Subtest Scores by Group for Session 1 (Upper Triangle) and Session 2 (Lower Triangle)

Subtest	PC	NO	CS	AS	MC	EI
Group 1 (N=171 Session 1[a]; N=332 Session 2)						
PC	—	.36	-.06	.08	.30	.32
NO	.14	—	.27	-.05	.19	.11
CS	.13	.56	—	.08	.18	.10
AS	.25	-.05	.01	—	.59	.49
MC	.32	.00	.08	.52	—	.56
EI	.34	-.02	.00	.56	.59	—
Group 2 (N=332 Session 1; N=167 Session 2[a])						
PC	—	.20	.22	.16	.25	.28
NO	.15	—	.62	-.04	.12	-.02
CS	.27	.41	—	.11	.21	.10
AS	.24	-.17	.01	—	.54	.57
MC	.34	-.01	.11	.56	—	.55
EI	.40	-.09	.00	.52	.54	—
Group 3 (N=333, both sessions)						
PC	—	.39	.34	.21	.36	.35
NO	.26	—	.63	-.03	.16	.07
CS	.27	.57	—	.02	.21	.09
AS	.22	-.03	.05	—	.53	.55
MC	.41	.18	.28	.52	—	.52
EI	.37	.11	.16	.58	.57	—

Note. Only factors with eigenvalues greater than zero are presented.

[a]Only computer Modes 1 and 3 for PC, and Mode 1 for NO and CS, were included since Mode 2 for PC and Mode 2 for NO and CS were found to be significantly different in the MANOVA analysis.

Factor 1 for cells C22 (.488) and P32 (.521), crossing both Session and Medium.

The second factor is a little cleaner, with its highest loadings found on the speeded subtests (NO and CS) in every cell except C11, where PC and AS have stronger loadings than CS (-.432 versus .350). In Group 1, Session 1, both PC and AS have stronger loadings than CS (.357 and -.347 versus .184). In Group 3 there is also a substantial loading on Factor 2 for subtest AS in each session (-.443 and -.385).

Assuming that a third factor is necessary to adequately explain the data

Table 18. Eigenvalues and Percentage of Common Variance Accounted for from Principal Factor Analyses of Subtest Scores by Group and Session

Group and factor	Session 1 Eigenvalue	Session 1 % of variance	Session 2 Eigenvalue	Session 2 % of variance
Group 1:	Computer (N=171)[a]		Paper-and-Pencil (N=332)	
1	1.740	67.8	1.719	65.5
2	.572	22.3	.906	34.5
3	.255	9.9		
Group 2:	Paper-and-Pencil (N=332)		Computer (N=167)[a]	
1	1.740	63.4	1.732	69.3
2	1.006	36.6	.768	30.7
Group 3:	Paper-and-Pencil (N=333)		Paper-and-Pencil (N=333)	
1	1.885	64.4	1.956	69.8
2	1.043	35.6	.846	30.2

Note. Only factors with eigenvalues greater than zero are presented.

[a]Only computer Modes 1 and 3 for PC, and Mode 1 for NO and CS, were included since Mode 2 for PC and Mode 2 for NO and CS were found to be significantly different in the MANOVA analysis.

in cell C11, it is defined by low to moderate loadings on PC (-.289) and CS (.389). The communality estimate for CS in cell C11 (.220) is lower than in any other cell, the next lowest being .337 in the other computer-administered cell C22.

Table 20 summarizes the covariance structure and confirmatory factor analysis results. The first model tested was the equality of the across-subtest covariance matrices within sessions. In both sessions this model was rejected as not fitting the data, with $\chi^2 = 104.57$ and $\chi^2 = 75.47$ with 42 degrees of freedom. A check on the sensitivity of the procedure was run testing the equality of the across-subtest covariance matrices for only the paper-and-pencil administrations within each session. This model was not rejected in either session ($\chi^2 = 22.42$ and $\chi^2 = 29.91$, with 21 degrees of freedom each), indicating that the covariance matrices for computer administration sessions were not equal to those from paper-and-pencil testing sessions.

The next model tested a less-strong equality that the factor pattern was invariant for each group within sessions. The principal factors analyses suggested that either two or three factors were present; thus, the first tests assumed the presence of three correlated factors for each group. Assuming that this model was not rejected, subsequent models would test the equality of the factor loadings, error variances and covariances, and factor covariances,

Table 19. Factor Loadings and Communalities (h^2) from Principal Factor Analyses of Subtest Scores by Group and Session

Group and subtest	Session 1 Factor 1	Factor 2	Factor 3	h^2	Group and subtest	Session 2 Factor 1	Factor 2	Factor 3	h^2
Group 1 (N=171): Computer					Group 1 (N=332): Paper-and-Pencil				
PC	.398	.357	-.289	.370	PC	.431	-.141	-	.205
NO	.256	.531	.092	.355	NO	.050	-.663	-	.442
CS	.188	.184	.389	.220	CS	.105	-.653	-	.437
AS	.634	-.347	.059	.526	AS	.665	.105	-	.453
MC	.778	-.054	.048	.611	MC	.720	.029	-	.519
EI	.688	-.073	-.081	.486	EI	.748	.095	-	.568
Group 2 (N=332): Paper-and-Pencil					Group 2 (N=167): Computer				
PC	.381	-.119	-	.159	PC	.488	-.299	-	.327
NO	.283	-.668	-	.526	NO	-.041	-.562	-	.317
CS	.409	-.596	-	.523	CS	.133	-.565	-	.337
AS	.639	.296	-	.496	AS	.673	.186	-	.487
MC	.696	.134	-	.502	MC	.715	.007	-	.512
EI	.674	.290	-	.430	EI	.714	.095	-	.519
Group 3 (N=333): Paper-and-Pencil					Group 3 (N=333): Paper-and-Pencil				
PC	.572	.155	-	.351	PC	.521	.107	-	.283
NO	.455	.591	-	.577	NO	.352	.571	-	.451
CS	.473	.544	-	.599	CS	.431	.533	-	.469
AS	.542	-.443	-	.490	AS	.589	-.385	-	.495
MC	.670	-.238	-	.506	MC	.727	-.113	-	.541
EI	.619	-.348	-	.505	EI	.707	-.251	-	.563

Note. Only factors with eigenvalues greater than zero are presented.

Table 20. Summary Table for Testing Models of Factor Equality Within Session, Showing Model, Chi-Square Value (χ^2), Degrees of Freedom (df), Goodness of Fit Index (GFI), Root-Mean Square Residual (RMR), and Estimated Probability (p)

	Hypothesis	χ^2	df	GFI	RMR	p<
	$\Sigma_{11} = \Sigma_{21} = \Sigma_{31}$	104.57	42	.977	2.206	.001**
	$\Sigma_{21} = \Sigma_{31}$	22.42	21	.989	2.090	.374
	$\Sigma_{12} = \Sigma_{22} = \Sigma_{32}$	75.47	42	.985	3.513	.001**
	$\Sigma_{12} = \Sigma_{32}$	29.91	21	.984	4.632	.094
A.	$\Lambda_{x_{11}} = \Lambda_{x_{21}} = \Lambda_{x_{31}}$	$\Theta\delta$ not positive definite for C11				
A.	$\Lambda_{x_{12}} = \Lambda_{x_{22}} = \Lambda_{x_{32}}$	56.58	21	.967	2.297	.001**
B.	$\Lambda_{x_{11}} = \Lambda_{x_{21}} = \Lambda_{x_{31}}$	Φ not positive definite for C22				
B.	$\Lambda_{x_{12}} = \Lambda_{x_{22}} = \Lambda_{x_{32}}$	Φ, $\Theta\delta$ not positive definite for C11				
C.	$\Lambda_{x_{11}} = \Lambda_{x_{21}} = \Lambda_{x_{31}}$	157.54	24	.940	2.239	.001**

**Statistically significant at $p < .01$.

Note. A is a model of the factor pattern:

$$\Lambda_x = \begin{bmatrix} 0 & 0 & X \\ 0 & X & 0 \\ 0 & X & 0 \\ X & 0 & 0 \\ X & 0 & 0 \\ X & 0 & 0 \end{bmatrix}$$

B is a model of the factor pattern:

$$\Lambda_x = \begin{bmatrix} 0 & 0 & X \\ 0 & X & 0 \\ 0 & X & 0 \\ X & 0 & 0 \\ X & 0 & X \\ X & 0 & X \end{bmatrix}$$

C is a model of the factor pattern:

$$\Lambda_x = \begin{bmatrix} 0 & X \\ 0 & X \\ 0 & X \\ X & 0 \\ X & 0 \\ X & 0 \end{bmatrix}$$

providing an indication of the degree of equality of the factor structures.

Only for Session 2 did LISREL converge to meaningful parameter estimates while testing a three-factor model. This model (Factor 1 = AS, MC, EI; Factor 2 = NO, CS; Factor 3 = PC; zeroes elsewhere) produced a significant chi-square (χ^2 = 56.58, 21 degrees of freedom) indicating non-fit. In Session 1, the matrix theta-delta was not positive definite for Group 1 (C11), indicating that the model did not fit the data for that cell.

A second three-factor pattern was hypothesized, allowing two more loadings to vary (Factor 1 = AS, MC, EI; Factor 2 = NO, CS; Factor 3 = PC, MC, EI; zeroes elsewhere), which lessened the pressure to fit many zero loadings. For this model, LISREL did not converge to meaningful parameter estimates in either session. The Session 1 results found both matrices phi and theta-delta to be non-positive definite for Group 1 (C11). In Session 2, phi was not positive definite for cell C22 (Group 2). Again, the non-meaningful estimates of these matrices imply that the model did not fit the data within those cells.

Next, a two-factor pattern model (Factor 1 = AS, MC, EI; Factor 2 = PC, NO, CS; zeroes elsewhere) was tried for Session 1, in case three factors were not necessary to explain the data adequately. This model fit the data less well than the most restrictive covariance equality model (χ^2 = 104.57, df = 42), with a χ^2 value of 157.54 with 24 degrees of freedom. The implication of this test is that neither a two-factor model nor a three-factor model fit the Session 1 data. Thus, the reason for the non-fit of the three-factor model could not be the specification of too many factors.

Within Subtest

Table 21 shows eigenvalues and the percentage of common variance accounted for from the within-subtest principal factors analysis by Form (11, 12 and 13) and Medium of administration. Appendix A shows the factor loadings for the first five principal factors from the same analyses. Only non-speeded subtests AS, MC, and EI were factor analyzed, because it was necessary to combine the data from examinees taking both Versions A and B in order to obtain enough examinees to meet factor analysis requirements. Subtest PC contains different items in Versions A and B, making such a combination impossible. Due to the finding of significant Session effects in the MANOVA analysis, all analyses were performed within Session 1, with Group 1 (computer administration) being compared against Group 3 (paper-and-pencil administration).

There were no major trends across subtests or forms, or large factor structure differences for the subtests, under different media of administration conditions. The factor structures of each form of each subtest under both administration media were adequately described by one-factor solutions. The large number of eigenvalues greater than 1.0 found for each subtest is likely an indication of the small examinee-per-item ratio for each factor analysis (4:1 or 5:1). The first factors accounted for 43.8% (MC Form 13, computer-administered) to 59.4% (AS Form 13, computer-administered) of the common variance in each subtest. These first factors are more than twice as large as the second factors in every case, suggesting the interpretation of a one-factor solution. All subtests, however, showed a trend for the first factor to be larger (i.e., account for more common variance) under computer administration than for paper-and-pencil administration. These differences

Table 21. Eigenvalues (E) and Percentage of Common Variance Accounted for (%), by the First Five Principal Factors from Subtests AS, MC, and EI by Form and Administration Medium

Test and factor	Form 11				Form 12				Form 13			
	Computer		Paper-and pencil		Computer		Paper-and pencil		Computer		Paper-and pencil	
	E	%	E	%	E	%	E	%	E	%	E	%
AS												
1	4.38	54.8	3.79	50.4	5.04	56.6	4.31	55.5	5.05	59.4	4.10	52.7
2	1.27	15.9	1.14	15.1	1.42	15.9	1.12	14.4	1.03	12.0	1.14	14.7
3	0.86	10.7	0.99	13.2	1.12	12.5	0.97	12.5	0.95	11.2	0.91	11.6
4	0.75	9.4	0.91	12.1	0.69	7.8	0.72	9.3	0.79	9.3	0.88	11.4
5	0.74	9.2	0.69	9.2	0.64	7.1	0.64	8.2	0.69	8.1	0.74	9.6
MC												
1	4.39	56.8	3.11	45.5	3.50	54.0	3.87	52.0	2.74	43.8	3.11	45.5
2	0.92	11.9	1.15	16.8	0.85	13.1	1.15	15.5	1.04	16.6	1.19	17.3
3	0.85	11.0	0.97	14.3	0.78	12.2	0.89	11.9	0.92	14.8	0.95	13.9
4	0.81	10.5	0.86	12.6	0.71	11.0	0.80	10.8	0.82	13.2	0.81	11.9
5	0.75	9.8	0.73	10.7	0.62	9.7	0.73	9.8	0.73	11.7	0.78	11.4
EI												
1	3.06	50.4	2.93	50.4	3.29	57.5	3.44	55.0	3.09	53.2	2.43	47.4
2	1.08	17.8	0.87	14.9	0.81	14.3	0.92	14.7	1.04	17.8	0.86	16.8
3	0.77	12.7	0.76	13.2	0.63	10.9	0.76	12.2	0.72	12.3	0.70	13.7
4	0.65	10.7	0.68	11.7	0.51	8.9	0.62	9.8	0.53	9.0	0.67	13.2
5	0.51	8.4	0.57	9.8	0.48	8.4	0.52	8.3	0.44	7.6	0.45	8.9

ranged from 11.3% in MC Form 11 to 1.1% in AS Form 12. There was no difference in EI Form 11, and there was a 1.7% difference in the other direction for MC Form 13, but these were the only exceptions.

The only large difference in computer- versus paper-and-pencil-administered subtest factor structure occurred for subtest MC, Form 11. For computer administration, Factor 1 accounted for 56.8% of the common variance, but for the paper-and-pencil administration, Factor 1 captured only 45.5%, causing an 11.3% difference. This difference was not matched in Forms 12 or 13 of subtest MC; in fact, for Form 13, the paper-and-pencil Factor 1 accounted for more common variance than did the Factor 1 from the computer administration (45.5% versus 43.8%).

Item Analysis

Conventional item statistics (proportion correct, point-biserial, and biserial item-test correlations) for each item in each non-speeded subtest (PC, AS, MC and EI) by Version (A or B) and Medium (computer or paper-and-pencil) are shown in Appendix B. There were no statistical analyses performed at the item level to compare particular values of these statistics, due to the mostly nonsignificant and inconsequential differences found for the Medium effect in the MANOVA analyses.

The important question of possible differences in graphical content items across administration media was addressed partially through the subtest-level MANOVAs. The distribution of graphical items allowed the comparisons of subtest MC across administration media to substitute for an item-level analysis. In fact, every subtest-level analysis performed on subtest MC could be used as a method to compare graphical versus non-graphical content items. By comparing computer administration versus paper-and-pencil administration differences for MC versus AS and EI, information on graphical items was obtained. The results of these comparisons were that differences in subtest MC across administration media were not significantly greater than the computer versus paper-and-pencil differences exhibited by AS or EI. The within-subtest factor analyses also showed no greater differences in the factor structure of MC across administration media than were found for subtests AS and EI.

IV. DISCUSSION

The purpose of this study was to determine the extent of equivalence of measurement properties for a battery of ASVAB subtests under conditions of computer versus paper-and-pencil administration. The subtests selected for study--PC, NO, CS, AS, MC, and EI--were those that presented particular problems for computer presentation.

The paper-and-pencil baseline analysis was performed to yield a lower bound against which computer versus paper-and-pencil differences could be judged, both within and between examinees. The only statistically significant effects found were those for Form on the AS subtest and Session for the NO, CS, and AS subtests. However, the statistically significant mean differences observed were small and not psychometrically meaningful, thus providing an acceptable basis for judging comparisons between computer and paper-and-pencil conditions.

Paragraph Comprehension Subtest

The PC subtest, chosen due to its paragraph-length items and multiple items per paragraph, showed differences in measurement properties across the medium of presentation and within the three modes of computer administration, with one Mode (CRT-2) differing significantly from the other two (CRT-1, CRT-3). The first Mode (CRT-1) displayed the paragraph in one scrolling field while the items appeared sequentially in a separate nonscrolling field beneath it. Each item was erased before the next one appeared, and the examinee, unable to retrieve the item, could therefore proceed only in a forward direction through the test. The second computer administration Mode (CRT-2) contained each paragraph and all relevant questions in a single scrolling field. In this mode, the entire screen was available for viewing the paragraph, with questions appearing after the final line of text as the paragraph was scrolled up the screen, allowing the examinee to move back and forth within each paragraph. For Sessions 1 and 2, this condition resulted in mean scores significantly lower than both the paper-and-pencil condition and the other two CRT conditions, suggesting examinee confusion or disorientation arising from this particular screen format. Another possibility is that because whole paragraphs and corresponding items never appeared on the screen simultaneously, a memory component was introduced and became more important in this condition than in the others. The final computer condition (CRT-3) contained separate scrolling fields for both the paragraphs and their related items, and provided an answer-sheet type of display at the top of the screen, allowing examinees to monitor their progress through the test. This condition also enabled examinees to return to any paragraph and to change their response to any item at any time during the test. In both sessions, this condition resulted in mean scores almost identical to those for the CRT-1 condition.

Clear differences were demonstrated within Session 1 between all three CRT conditions and the paper-and-pencil administration condition. In Session 2 however, equivalent scores were obtained for the paper-and-pencil, CRT-1, and CRT-3 conditions, with the mean paper-and-pencil score being fairly constant, while the scores for all three CRT groups were significantly larger compared to Session 1. Only the CRT-2 condition still yielded significantly lower scores than paper-and-pencil in Session 2. This finding suggests that those who took the paper-and-pencil PC test first, followed by the computer PC test, may have benefitted from practice effects in the second session, implying that a more extensive practice sequence and perhaps a more detailed instruction set preceding the administration of the computerized PC test may be appropriate. In light of the absence of comparable differences on the other subtests, another possibility may be that, due to the combination of lack of familiarity with the computer medium and the complexity of the PC subtest, the first subtest administered in the sequence produced a heightened level of anxiety which attenuated test scores for those tested by computer in the first session. The equivalence of the CRT-1 and CRT-3 conditions may be due to a tendency for examinees to ignore the more complex features of the CRT-3 condition, such as the ability to return to earlier paragraphs, responding instead in a manner similar to that of the CRT-1 format, i.e., proceeding straight through the test item by item without backtracking. The lower performance for CRT-2 may have resulted from the increased memory requirement caused by the particular screen configuration used.

The reliability analyses found that the PC subtest obtained lower inter-

nal consistencies of a meaningful magnitude (> .10) when computer-administered. However, the test-retest reliabilities when one administration was by computer were not significantly lower than the paper-and-pencil-only test-retest reliabilities. In fact, the Group 2 test-retest correlations were significantly higher (for Modes 1 and 3) than those from the paper-and-pencil-only group.

Speeded Subtests

The NO and CS subtests, chosen on the basis of their speeded nature, were administered by paper-and-pencil and by two CRT conditions. The first computer mode presented one item at a time on the screen, with each item being erased and replaced by the next item following the examinee's response. The second mode filled the screen with a block of items, with the examinee responding to the first item in the block. This item was then erased and the remaining items in the block shifted upward following the examinee's response. The number of items present on the screen decreased in this manner with each response until, following the response to the last item in the block, a new block of items appeared on the screen.

Although the initial analysis revealed significant differences between the paper-and-pencil and CRT media for both the NO and CS subtests in Session 1, further evaluation demonstrated equivalence between one CRT condition and the paper-and-pencil medium in both sessions. For the NO subtest, this equivalence was found between the CRT-1 single-item mode and the paper-and-pencil condition, with the CRT-2 group scoring significantly lower. For CS, the equivalent conditions were the CRT-2 multiple-item screen mode and the paper-and-pencil condition, with the CRT-1 single-item group scoring much higher than either of the others. It should be noted that for every comparison between the CS CRT-2 and paper-and-pencil conditions, a consistent trend of marginally lower CRT scores was observed, arguing against absolute equivalence.

For both speeded subtests, a consistent relationship was found between the two computer conditions, with the single-item presentation causing higher scores than the multiple-item presentation in every case, the multiple-item mode perhaps causing distraction or confusion for the examinees. For the NO subtest, this resulted in highly attenuated scores for the CRT-2 condition as compared with paper-and-pencil results. For the CS subtest, the single-item computer mode provided such a marked increase in performance over paper-and-pencil administration that the attenuating effects of multiple-item computer presentation brought scores more in line with those for the paper-and-pencil medium. This finding may be deceptive, however. The CS single-item presentation condition may actually be more parallel to the paper-and-pencil performance, differing only by a scaling factor due to the greater response speed afforded the CRT examinee. The multiple-item response condition may be subject to negative effects arising from rapid response times which are then negated by the distracting influence of the upward-shifting multiple-item screen format, as found with the NO subtest.

The test-retest reliability analysis showed for both NO and CS that administration by computer in one session significantly lowered the correlation of computer-administered scores with those from paper-and-pencil administration, in comparison to the paper-and-pencil baseline analysis. This suggests that the administration of speeded tests by computer might result in test

scores with somewhat different score properties than the same speeded tests administered by paper-and-pencil.

Graphical Subtests

The AS, MC, and EI subtests, chosen for their graphical and standard multiple-choice text content, were presented in a single CRT mode in addition to paper-and-pencil presentation. The format of the computer presentation was identical to that of the paper-and-pencil tests with digitized graphical images, copied directly from the ASVAB test booklets, appearing on the computer screen with the appropriate text.

The results of the analyses for these subtests indicated a straightforward equivalence between the computer and paper-and-pencil media for all three subtests. The clarity of this finding can be attributed, at least in part, to the construction of finely detailed computer representations of all graphical images. A direct test of the equivalence of examinee perception of these images was provided by the comparison between CRT and paper-and-pencil conditions within Session 1 for the MC subtest, the content of which is almost entirely graphical. This comparison identified no differences between the two modes of presentation, thus suggesting equivalence.

Subtest Structure

The interrelationships of this battery of subtests as a whole were somewhat different under the conditions of different administration methods. The principal factor analyses suggested that the factor structure of the subtest scores for computer-administered Group 1 Session 1 contained three factors, whereas only two factors were needed to explain the data from the other first session paper-and-pencil administrations. The LISREL confirmatory analysis supported these results by rejecting a model of the equality of the subtest covariance matrices when a computer administration group was included, but not rejecting a model of equal covariance matrices when only covariance matrices from the paper-and-pencil groups were included. All of the less stringent tests of the equality of factor structure models across administration media found either a rejection of the model or the presence of non-positive definite covariance matrices. A finding of non-positive definite covariance matrices is an indication that the model did not provide a suitable fit to the data.

These negative LISREL findings could be due to the slight factor structure difference found in the principal factors analyses or to the extreme power of the LISREL procedure when sample sizes and degrees of freedom are large. The model of equal factor structure that did converge to meaningful parameter estimates produced a significant chi-square, but the degree of nonfit was quite small. The estimated covariance matrices for Groups 1, 2, and 3 produced only zero, one, and two significant normalized residuals, respectively. Thus, although the model did obtain a significant chi-square value, it should not be rejected entirely.

V. CONCLUSIONS

The results of the present study suggest that attaining equivalent test results between computer and paper-and-pencil administrations is feasible for

the power subtests of the ASVAB. For the case of standard multiple-choice text items and those with graphical content, all that may be required is careful development of testing software, with adequate attention to the clarity and detail of graphical images.

For the speeded subtests, the present findings indicate that the mode of screen presentation of test items can drastically influence the level of examinee performance. For the NO subtest, an indication of equivalence between the single-item CRT and paper-and-pencil presentation conditions was established. For CS, however, although the multiple-item screen condition produced similar scores to those from the paper-and-pencil medium, the trend over multiple comparisons between the two showed that the computer scores were marginally lower in every case. This finding suggests that further research on alternative screen configurations and enlarged computer instruction and practice sets may be appropriate to bring computer performance more in line with its paper-and-pencil counterpart. Continued investigation may also reveal that the higher-scoring single-item screen condition is actually more consistent with paper-and-pencil performance, and that a scaling factor is required to compensate for the faster response time made possible by the substitution of a computer keyboard for a paper-and-pencil answer sheet.

The PC subtest offers perhaps the greatest challenge to equivalence in the ASVAB battery. For this subtest, the present results demonstrate the sensitivity of examinee performance to alternative screen presentation modes for this rather complex test. The findings support, however, the benefits of pre-test practice, simplicity of screen format design, and detailed instruction sets in equalizing computer and paper-and-pencil performance. Recommendations for further research include experimentation with varying numbers of practice items, alternative screen formats, instruction sets emphasizing specific aspects of the computerized PC subtest, and administration of the PC subtest after administration of other subtests in the ASVAB battery to allow greater examinee familiarity with the computer before the presentation of the PC items.

REFERENCES

Cooley, W., & Lohnes, P. (1971). Multivariate data analysis. New York: Wiley.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. Psychometrika, 16, 297-334.

Feldt, L. S. (1980). A test of the hypothesis that Cronbach's alpha reliability coefficient is the same for two tests administered to the same sample. Psychometrika, 45, 99-106.

Fisher, R. A. (1921). On the "probable error" of a coefficient of correlation deduced from a small sample. Metron, 1, 3-32.

Green, B. F., Bock, R. D., Humphreys, L. G., Linn, R. L., & Reckase, M. D. (1982, May). Evaluation plan for the computerized adaptive vocational aptitude battery (Research Report 82-1). Baltimore, MD: The Johns Hopkins University, Department of Psychology.

Hakstian, A. R., & Whalen, T. E. (1976). A k-sample significance test for independent alpha coefficients. Psychometrika, 41, 219-231.

Harman, H. H. (1976). Modern factor analysis (3rd ed.). Chicago: University of Chicago Press.

Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. Psychometrika, 36, 409-426.

Jöreskog, K. G., & Sörbom, D. (1984). LISREL VI: Analysis of linear structural relationships by the method of maximum likelihood. Mooresville, IN: Scientific Software, Inc.

Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. Psychometrika, 2, 151-160.

Morrison, D. F. (1976). Multivariate statistical methods (2nd. ed.). New York: McGraw-Hill Co.

Neter, J. & Wasserman, W. (1974). Applied linear statistical models. Homewood, IL: Irwin.

O'Brien, R. G., & Kaiser, M. K. (1985). MANOVA method for analyzing repeated measures designs: An extensive primer. Psychological Bulletin, 97, 316-333.

Olson, C. L. (1976). On choosing a test statistic in multivariate analysis of variance. Psychological Bulletin, 83, 579-586.

Prestwood, J. S., Vale, C. D., Massey, R. H., & Welsh, J. R. (1985, September) Armed Services Vocational Aptitude Battery: Development of Forms 11, 12, and 13 (AFHRL-TR-85-16(I), AD-A160584). Brooks Air Force Base, TX: Air Force Human Resources Laboratory, Manpower and Personnel Division.

Winer, B. J. (1971). Statistical principles in experimental design (2nd. ed.). New York: McGraw-Hill.

Wood, R. L., & Lord, F. M. (1976). A user's guide to LOGIST (Research Memorandum 76-4). Princeton, NJ: Educational Testing Service.

Wood, R. L., Wingersky, M. S., & Lord, F. M. (1976). LOGIST--A computer program for estimating examinee ability and item characteristic curve parameters (Research Memorandum 76-6). Princeton, NJ: Educational Testing Service.

APPENDIX A: ITEM FACTOR LOADINGS AND COMMUNALITIES BY SUBTEST AND FORM FOR TWO PRESENTATION MEDIA

Table A-1. Factor Loadings and Communalities (h^2) for the First Five Principal Factors from Session 1 Subtest AS Form 11 for Group 1 (Computer) and Group 3 (Paper-and-Pencil)

	Computer						Paper-and-Pencil					
Item	1	2	3	4	5	h^2	1	2	3	4	5	h^2
1	.491	-.389	-.138	-.082	.301	.509	.025	-.066	.257	.354	.072	.202
2	.275	.170	-.250	-.243	-.054	.229	.235	.081	.145	-.227	.001	.135
3	.329	-.352	-.206	-.175	.218	.353	.417	.367	-.033	.163	-.062	.340
4	.445	.231	-.005	.035	-.072	.259	.344	.193	.098	-.048	.482	.401
5	.364	-.149	.004	-.203	-.337	.310	.454	-.270	.049	-.465	.019	.499
6	.252	.342	-.134	.264	.207	.312	.114	-.009	-.085	-.039	.126	.038
7	.374	.201	.174	-.266	.096	.292	.389	.220	-.125	.107	-.214	.273
8	.408	-.046	-.128	.303	-.060	.280	.345	.214	-.545	-.049	-.060	.468
9	.469	.087	-.181	-.218	.021	.309	.354	.185	.061	-.394	-.244	.378
10	.569	.153	-.229	.084	.072	.413	.309	-.301	-.045	-.280	.074	.273
11	.534	-.002	.154	.019	-.104	.320	.512	-.097	.015	.300	-.135	.381
12	.384	.249	.220	-.228	.139	.330	.430	-.187	.344	.030	-.275	.415
13	.562	.098	-.259	.028	-.248	.456	.414	-.330	-.314	.135	-.150	.420
14	.131	-.104	-.085	.097	-.415	.217	.136	.388	.147	.126	-.099	.217
15	.379	-.175	.107	.164	.185	.247	.397	-.279	-.258	.139	.149	.344
16	.349	-.279	.295	-.036	-.028	.289	.607	-.033	-.075	-.041	-.028	.378
17	.331	.188	.010	-.103	.044	.158	.442	.241	-.085	.014	.253	.325
18	.351	.389	.115	.158	-.166	.340	.229	-.236	.075	.196	.102	.163
19	.601	-.231	.184	-.102	-.069	.465	.607	.160	.150	-.041	.071	.424
20	.155	.161	.341	.035	.023	.168	.292	-.227	.312	.087	.141	.262
21	.692	-.228	.171	.062	-.153	.589	.589	.090	.273	-.033	.008	.431
22	.309	-.196	-.148	.331	.050	.268	.432	-.230	-.154	.123	.033	.280
23	.417	.358	.067	.052	.178	.341	.257	.133	-.087	.093	.151	.123
24	.394	-.144	.207	.233	.137	.292	.445	.023	.091	.096	-.183	.250
25	.383	.024	-.265	-.096	.042	.229	.240	.076	-.069	-.055	.109	.083

Table A-2. Factor Loadings and Communalities (h^2) for the First Five Principal Factors from Session 1 Subtest AS Form 12 for Group 1 (Computer) and Group 3 (Paper-and-Pencil)

	Computer						Paper-and-Pencil					
Item	1	2	3	4	5	h^2	1	2	3	4	5	h^2
1	.524	-.550	.085	.045	-.000	.588	.486	.064	-.091	-.127	-.054	.268
2	.501	-.165	.025	-.015	-.138	.298	.322	.292	.174	-.063	.231	.277
3	.246	-.430	.028	.205	.203	.33	.251	.175	.339	-.060	.165	.239
4	.467	-.513	.004	-.035	-.110	.496	.386	.022	-.265	.115	-.111	.246
5	.536	-.028	-.307	-.054	.162	.413	.304	.000	-.065	-.293	-.149	.205
6	.316	-.053	.258	-.220	-.104	.229	.410	.219	.014	-.023	-.132	.234
7	.318	.269	.464	-.055	.004	.393	.639	-.027	-.076	.055	-.223	.468
8	.385	.121	.444	.114	.203	.415	.428	.377	-.204	-.056	.026	.371
9	.624	-.008	-.070	-.169	-.094	.432	.390	-.138	.114	-.068	.001	.189
10	.511	.140	-.161	-.205	.187	.384	.406	.140	.225	.271	.110	.321
11	.691	-.177	.036	-.162	.170	.567	.643	.079	.041	-.119	-.278	.514
12	.422	-.122	.415	-.048	-.142	.388	.438	.174	-.046	-.040	-.058	.229
13	.468	.194	-.194	-.351	.124	.433	.451	-.344	-.150	.040	.116	.360
14	.166	.000	.053	.164	.354	.183	.475	.308	.182	-.008	.014	.355
15	.502	.095	-.165	.132	-.018	.306	.377	.000	-.036	.279	.256	.287
16	.288	.340	.220	.221	.140	.317	.376	.164	-.381	-.223	.348	.486
17	.441	.286	-.152	.027	-.158	.326	.274	-.238	-.350	-.122	-.075	.275
18	.440	.098	.019	.356	-.112	.343	.372	-.213	.080	-.235	.189	.282
19	.588	.014	.047	-.069	-.167	.382	.481	-.334	.106	.054	-.023	.358
20	.305	.290	-.084	.084	-.197	.230	.380	-.303	.243	-.031	.155	.321
21	.071	-.013	-.161	-.039	.262	.102	.240	.160	-.113	.146	-.055	.120
22	.516	.250	-.180	.036	-.076	.369	.451	-.339	.119	-.170	.097	.371
23	.392	.040	-.017	.091	-.164	.191	.370	.031	-.165	.316	.130	.282
24	.475	-.096	-.331	.330	-.030	.455	.378	-.204	-.088	.376	-.098	.344
25	.475	.258	.102	.018	.149	.325	.359	.012	.397	.017	-.230	.340

Table A-3. Factor Loadings and Communalities (h^2) for the First Five Principal Factors from Session 1 Subtest AS Form 13 for Group 1 (Computer) and Group 3 (Paper-and-Pencil)

	Computer						Paper-and-Pencil					
Item	1	2	3	4	5	h^2	1	2	3	4	5	h^2
1	.481	-.126	-.141	.004	.208	.311	.316	-.437	.298	-.129	.088	.405
2	.250	.481	-.026	.426	.052	.480	.328	.260	.193	.172	-.141	.263
3	.436	-.312	-.280	-.098	.289	.460	.396	-.240	.033	.156	-.319	.342
4	.385	-.161	.491	-.087	.091	.432	.241	.412	.045	-.112	.237	.299
5	.263	.251	-.172	.091	-.086	.178	.239	.118	-.152	-.237	.082	.157
6	.302	.020	.298	.232	.003	.234	.287	.398	.332	-.041	-.037	.354
7	.105	-.025	-.336	.079	.065	.135	.148	.346	-.119	-.012	.147	.178
8	.346	.387	.162	-.224	.212	.392	.510	.124	-.119	-.107	-.201	.342
9	.356	-.137	.279	.317	-.136	.343	.352	-.335	.234	-.036	.318	.394
10	.562	.218	-.057	.038	.044	.370	.433	-.132	-.227	-.134	-.089	.283
11	.584	-.235	.188	.117	.185	.480	.439	.197	-.052	-.391	-.140	.408
12	.290	.155	-.206	-.011	.073	.156	.235	.072	.319	.008	.123	.178
13	.375	-.075	.031	-.030	-.030	.149	.429	.020	.259	-.187	-.214	.333
14	.393	.235	.150	-.387	.171	.411	.385	.151	-.055	.174	.083	.212
15	.602	-.036	-.023	.073	-.306	.464	.767	-.198	.107	.010	.108	.651
16	.206	.077	-.093	-.065	-.109	.073	.297	.020	.091	.504	.061	.355
17	.602	-.216	-.058	-.127	-.087	.437	.596	-.046	.069	-.036	-.072	.369
18	.477	.339	.123	-.119	-.054	.375	.401	.062	-.175	-.104	.332	.317
19	.622	-.120	-.281	.153	.123	.519	.470	-.064	.011	-.104	-.004	.236
20	.641	.024	-.108	-.169	-.118	.466	.450	.118	-.263	.211	-.045	.333
21	.325	-.140	.158	.216	.242	.256	.140	.014	.062	.044	-.308	.120
22	.434	.054	.054	-.004	.113	.208	.293	.150	.038	.375	-.031	.252
23	.477	-.155	.056	-.243	-.333	.426	.541	.006	-.084	-.022	-.028	.301
24	.595	.036	-.136	.037	-.065	.380	.329	-.222	-.385	.025	-.099	.316
25	.539	-.004	-.023	.089	-.261	.367	.486	-.114	-.215	.155	.225	.371

Table A-4. Factor Loadings and Communalities (h^2) for the First Five Principal Factors from Session 1 Subtest MC Form 11 for Group 1 (Computer) and Group 3 (Paper-and-Pencil)

	Computer						Paper-and-Pencil					
Item	1	2	3	4	5	h^2	1	2	3	4	5	h^2
1	.491	-.026	.115	-.011	-.111	.268	.279	-.114	.213	-.100	.104	.157
2	.178	.490	-.122	.110	-.021	.300	.196	.080	.037	.324	-.113	.164
3	.586	-.198	.019	-.150	-.203	.448	.138	-.397	.258	-.159	.035	.270
4	.293	-.056	-.217	-.173	.047	.169	.280	.028	.057	.180	-.415	.287
5	.328	.095	-.051	-.180	.120	.167	.303	.542	.238	.122	.107	.470
6	.256	.006	-.022	.096	-.273	.150	.227	.369	.037	-.015	.018	.190
7	.508	.048	-.229	.213	-.288	.443	.163	.034	.177	.015	.014	.059
8	.444	.028	.229	.103	-.096	.270	.385	-.085	.229	.156	.026	.233
9	.496	-.021	-.101	-.369	.036	.394	.246	-.016	-.122	-.308	.131	.188
10	.508	-.062	.182	.065	-.018	.300	.478	-.001	-.143	-.085	.244	.316
11	.464	-.216	-.106	-.219	.241	.380	.437	.063	.088	-.113	-.224	.266
12	.511	-.052	.107	-.027	.069	.281	.366	-.144	-.301	-.217	-.165	.320
13	.335	-.420	.015	.131	.085	.314	.499	-.019	-.046	.422	.200	.470
14	.529	-.318	.034	.074	-.094	.397	.306	-.030	-.061	.085	.243	.165
15	.450	.217	.201	-.151	.137	.332	.433	-.131	-.177	-.115	.072	.255
16	.328	.141	.276	-.059	-.262	.276	.333	.196	.336	-.173	.097	.303
17	.361	-.076	.196	.321	.039	.280	.468	.302	-.082	-.182	-.022	.351
18	.273	.276	-.259	-.027	-.171	.248	.274	-.129	.416	.068	-.047	.272
19	.520	-.069	-.419	.169	-.005	.480	.491	-.301	-.036	-.005	-.268	.405
20	.427	.274	.090	-.116	-.080	.286	.256	-.027	-.238	.015	.243	.182
21	.378	.117	-.248	.073	.334	.336	.529	.215	-.085	-.248	-.255	.460
22	.423	.137	.027	.424	.173	.409	.546	-.356	-.094	.294	-.000	.521
23	.465	.041	-.138	-.151	-.005	.260	.209	.112	-.319	.131	.085	.182
24	.285	.101	.128	.156	.396	.289	.100	.221	-.218	.143	-.133	.145
25	.306	.138	.304	-.172	.105	.246	.284	-.120	.170	-.176	.174	.186

Table A-5. Factor Loadings and Communalities (h^2) for the First Five Principal Factors from Session 1 Subtest MC Form 12 for Group 1 (Computer) and Group 3 (Paper-and-Pencil)

	Computer						Paper-and-Pencil					
Item	1	2	3	4	5	h^2	1	2	3	4	5	h^2
1	.188	-.308	-.098	-.050	.054	.145	.097	-.247	-.079	.241	-.142	.155
2	.275	-.036	.029	-.320	.087	.188	.259	-.080	.241	.035	.227	.185
3	.342	.293	-.337	.025	-.293	.403	.425	.056	-.253	-.195	-.053	.289
4	.190	-.027	-.217	.241	-.090	.150	.382	.303	.304	.112	-.217	.391
5	.363	-.131	.011	.416	-.020	.323	.402	-.129	.168	.073	.179	.245
6	.117	-.197	.131	-.161	.062	.100	.220	.275	.489	.144	.155	.409
7	.294	.109	.374	-.056	-.169	.270	.402	-.083	.188	.148	-.321	.329
8	.359	-.008	-.007	.085	.105	.147	.165	-.018	-.138	.003	.243	.106
9	.503	-.175	-.196	-.021	.142	.343	.551	-.074	.003	.115	-.203	.364
10	.341	.353	.199	-.000	.083	.288	.328	.518	-.132	.136	-.094	.422
11	.455	.165	.017	-.022	-.051	.238	.502	-.100	.013	-.373	-.127	.418
12	.498	.105	.076	-.238	-.233	.376	.592	-.039	.231	-.218	.104	.465
13	.146	.255	-.192	.116	.149	.159	.472	-.179	-.032	.167	.047	.286
14	.345	.032	.191	.148	-.052	.181	.400	.109	-.235	-.068	.182	.265
15	.184	.144	.308	.217	.317	.298	.414	-.285	-.031	.230	.018	.307
16	.451	.217	.033	-.020	-.087	.260	.301	-.275	-.005	.161	-.085	.200
17	.509	-.061	-.103	-.100	.159	.309	.410	-.105	.119	-.157	.165	.245
18	.473	-.157	-.247	.141	-.065	.334	.453	.230	-.315	-.014	-.123	.373
19	.359	-.104	.183	.323	-.127	.295	.247	.050	-.078	.270	.261	.211
20	.529	.207	-.134	-.094	.064	.354	.533	-.169	-.121	.160	-.080	.359
21	.438	-.266	-.068	-.057	-.004	.271	.287	.359	.031	-.105	.120	.238
22	.337	.130	-.133	-.139	.320	.271	.514	.283	-.101	-.016	-.127	.371
23	.167	-.104	.058	-.087	-.285	.131	.291	-.040	.136	-.282	-.025	.185
24	.480	-.193	.172	.024	.126	.314	.427	-.040	-.222	.049	.334	.348
25	.459	-.221	.155	-.150	-.100	.317	.270	-.304	.022	-.309	-.098	.271

Table A-6. Factor Loadings and Communalities (h^2) for the First Five Principal Factors from Session 1 Subtest MC Form 13 for Group 1 (Computer) and Group 3 (Paper-and-Pencil)

	Computer						Paper-and-Pencil					
Item	1	2	3	4	5	h^2	1	2	3	4	5	h^2
1	.402	-.360	-.103	-.146	-.123	.339	.412	.200	.091	-.096	-.272	.020
2	.331	.120	-.279	.087	-.057	.213	.086	-.131	-.196	.241	-.139	.410
3	.383	.203	.123	.074	.114	.222	.386	.122	.309	.044	-.058	.660
4	.598	-.248	.086	-.117	-.011	.440	.384	.503	-.077	-.044	-.106	.214
5	.426	.069	.163	-.229	-.253	.330	.520	-.313	-.118	.200	-.185	.573
6	.272	.147	-.295	.264	-.107	.265	.277	-.432	.033	-.228	.074	.224
7	.405	-.180	-.002	-.210	.276	.317	.258	-.167	-.198	-.467	.038	.549
8	.287	-.012	.065	.070	.213	.137	.325	-.296	.122	.188	.176	.750
9	.330	.274	-.082	-.029	.158	.217	.409	-.063	-.087	-.007	-.113	.919
10	.375	.187	-.084	-.080	-.084	.197	.552	-.361	.141	-.161	-.014	.820
11	.379	.024	-.432	-.076	-.015	.338	.428	.045	.051	-.002	.186	.232
12	.327	.040	-.328	.135	.190	.271	.111	.215	.130	-.199	-.328	.231
13	.126	-.075	.239	.127	-.060	.098	.196	.037	.544	.080	.070	.475
14	.208	-.192	.045	.153	.011	.106	.203	.275	.043	.024	.222	.690
15	.331	.316	.031	-.285	-.138	.311	.389	.167	.227	.159	.233	.113
16	.313	.111	.317	.170	-.095	.249	.253	.028	-.162	.233	.291	.303
17	.120	-.243	.198	-.039	.194	.152	.410	-.104	.237	-.033	-.121	.517
18	.241	.136	.137	.388	.138	.265	.386	.093	-.092	.155	-.161	.171
19	.194	.161	.086	.260	.208	.182	.060	.082	-.021	.061	.240	.726
20	.273	.163	.380	-.011	.022	.246	.574	.070	-.020	-.027	-.057	.396
21	.335	.104	.102	-.195	-.082	.179	.235	-.136	-.154	.343	-.103	.262
22	.235	-.370	-.040	.259	-.012	.262	.340	.300	-.214	-.049	.174	.847
23	.285	.023	.021	.276	-.477	.387	.207	.126	-.053	-.218	.136	.283
24	.383	.032	-.003	-.132	.242	.224	.240	-.083	-.099	-.207	.294	.039
25	.324	-.423	.008	.061	-.089	.296	.480	.085	-.403	.016	-.039	.024

Table A-7. Factor Loadings and Communalities (h^2) for the First Five Principal Factors from Session 1 Subtest EI Form 11 for Group 1 (Computer) and Group 3 (Paper-and-Pencil)

	Computer						Paper-and-Pencil					
Item	1	2	3	4	5	h^2	1	2	3	4	5	h^2
1	.459	.263	.015	.425	-.069	.466	.284	-.295	.031	-.087	.064	.180
2	.257	.016	-.084	-.238	-.033	.131	.208	-.060	-.379	.141	.053	.213
3	.229	.595	-.117	.107	.054	.435	.227	.185	.081	-.004	.426	.274
4	.095	-.216	-.420	.205	.249	.338	.220	-.139	.183	-.354	-.136	.246
5	-.000	.089	.264	.224	.233	.182	.226	.556	.140	-.195	-.063	.423
6	.482	.210	-.094	.138	-.162	.331	.464	-.019	.057	.372	-.049	.361
7	.476	-.296	-.022	.006	-.030	.316	.248	.313	.104	.133	-.324	.294
8	.485	.156	-.154	-.135	-.311	.399	.381	-.001	-.114	.246	.211	.264
9	.385	-.053	-.120	.143	-.162	.213	.359	-.081	-.122	.170	-.140	.199
10	.442	.204	.103	-.238	.033	.306	.509	-.094	.169	.202	-.052	.341
11	.507	.166	.063	-.087	.299	.386	.672	.111	-.384	-.109	.010	.624
12	.242	.133	.195	.090	-.022	.123	.338	.109	-.087	-.032	-.183	.168
13	.460	-.292	-.380	-.016	.094	.452	.383	-.166	.332	.080	-.041	.293
14	.512	-.198	.022	-.084	.050	.311	.543	-.092	.128	-.232	.087	.382
15	.366	-.295	.367	.025	-.057	.360	.193	-.002	.153	.191	.086	.105
16	.395	-.010	.206	-.049	.281	.280	.430	-.337	.057	-.103	-.052	.316
17	-.021	.127	.137	-.171	-.079	.071	.115	.238	.091	-.002	.320	.181
18	.311	-.344	.240	.174	-.193	.341	.321	-.064	.239	-.100	.136	.193
19	.542	-.068	.045	.069	.061	.309	.534	-.001	-.318	-.265	.008	.458
20	.456	.069	-.032	-.314	.034	.314	.450	.246	.104	.035	-.084	.282

Table A-8. Factor Loadings and Communalities (h^2) for the First Five Principal Factors from Session 1 Subtest EI Form 12 for Group 1 (Computer) and Group 3 (Paper-and-Pencil)

	Computer						Paper-and-Pencil					
Item	1	2	3	4	5	h^2	1	2	3	4	5	h^2
1	.353	-.196	.073	.097	.016	.178	.333	-.121	-.046	-.285	.048	.211
2	.333	.006	-.370	-.032	.018	.249	.249	.324	-.003	.103	-.101	.188
3	.295	.186	.162	.062	.031	.153	.380	-.255	-.020	-.203	-.169	.280
4	.257	-.397	.242	.074	.093	.297	.322	-.265	-.077	.241	.006	.239
5	.432	.049	-.057	.069	-.240	.255	.356	-.023	.037	-.132	-.145	.167
6	.421	.018	.138	-.224	.331	.357	.288	-.082	.132	.227	.402	.321
7	.446	-.151	.241	-.104	-.115	.305	.587	-.006	-.052	-.211	-.101	.403
8	.434	-.095	-.411	.010	.144	.388	.479	-.035	.218	.303	-.315	.470
9	.295	-.183	-.104	.351	.090	.263	.333	-.155	.095	-.029	.156	.169
10	.497	.086	.084	.314	.005	.361	.175	.201	.414	-.031	-.167	.272
11	.386	.196	.058	.000	.181	.224	.412	-.181	.120	.208	.019	.261
12	.498	-.416	-.075	-.049	.101	.439	.620	-.245	.183	-.013	-.039	.481
13	.444	.027	.088	.095	-.352	.339	.247	-.044	.071	-.053	.205	.113
14	.258	.367	.020	.228	.046	.256	.461	-.086	-.520	.110	.002	.503
15	.660	.155	-.074	-.021	.047	.469	.600	-.026	-.060	.166	.053	.395
16	.440	.143	.157	-.073	-.017	.245	.534	-.020	-.191	-.236	.024	.379
17	.361	-.234	.005	-.129	-.216	.249	.468	.355	.139	-.186	.232	.454
18	.497	.130	-.120	-.321	-.145	.403	.487	.287	.134	-.062	.037	.343
19	.232	.151	-.167	-.025	-.035	.107	.273	.370	-.095	.151	.092	.252
20	.300	.159	.201	-.109	.128	.185	.291	.374	-.301	.108	-.132	.345

Table A-9. Factor Loadings and Communalities (h^2) for the First Five Principal Factors from Session 1 Subtest EI Form 13 for Group 1 (Computer) and Group 3 (Paper-and-Pencil)

	Computer						Paper-and-Pencil					
Item	1	2	3	4	5	h^2	1	2	3	4	5	h^2
1	.428	-.035	-.173	-.245	-.136	.293	.234	.441	-.045	-.173	.018	.282
2	.542	-.005	-.283	.105	-.034	.387	.160	-.075	.084	-.091	.366	.181
3	.593	.086	.153	-.051	-.180	.418	.212	-.243	.099	.047	-.124	.132
4	.497	-.182	-.130	.237	.071	.359	.337	.403	-.193	.111	.042	.328
5	.436	.240	.105	.050	.030	.263	.478	-.048	-.162	-.119	-.006	.271
6	.536	-.199	-.302	.097	.122	.443	.608	-.067	-.110	-.088	-.044	.397
7	.547	.525	-.120	.058	.063	.598	.523	-.015	-.216	.019	.151	.344
8	.418	-.408	.161	.037	-.074	.375	.449	-.064	.004	.137	.035	.226
9	.445	-.330	.059	-.087	.085	.326	.261	.002	.048	-.454	-.085	.284
10	.277	.014	.337	.023	.116	.205	.429	-.324	-.167	.164	.078	.350
11	.194	.485	-.085	-.027	-.055	.284	.310	.092	.064	.184	-.158	.167
12	.426	-.136	-.084	-.153	-.047	.233	.256	-.053	.325	-.345	.119	.308
13	.361	.087	.152	-.120	-.053	.178	.375	-.111	.049	-.171	-.068	.190
14	.136	-.029	.132	-.225	-.143	.108	.382	-.158	.226	.055	-.260	.293
15	.270	.087	.447	.020	.277	.358	.309	.263	.137	-.006	-.263	.253
16	.257	.167	-.073	-.044	.282	.181	.382	-.155	-.035	.140	.146	.212
17	.328	-.090	.022	-.284	.040	.199	.351	.269	.067	.282	-.001	.280
18	.404	-.021	.149	.218	-.226	.285	.125	.300	.319	.057	.226	.263
19	.056	-.175	-.007	.225	.167	.112	.159	-.040	.404	.118	.086	.212
20	.049	.073	.142	.312	-.266	.196	-.035	-.114	.277	.207	.007	.134

APPENDIX B: ITEM STATISTICS (DIFFICULTY, POINT-BISERIAL, AND BISERIAL CORRELATION) BY SUBTEST, FORM, AND VERSION FOR TWO PRESENTATION MEDIA

Table B-1. Item Difficulties (p), Point-Biserial Correlations (rpb), and Biserial Correlations (rbis), for Version A of the PC Subtest by Session and Test Administration Medium

Session and item	Computer									Paper-and-Pencil								
	Mode 1			Mode 2			Mode 3			Mode 1[a]			Mode 2[a]			Mode 3[a]		
	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis
Session 1																		
1	1.00	0.000	0.000	1.00	0.000	0.000	1.00	0.000	0.000	.947	.354	.737	.966	.230	.549	.966	-.033	-.079
2	.879	.135	.219	.824	.266	.392	.912	.272	.482	.860	.202	.315	.828	.013	.020	.948	.127	.266
3	.966	-.003	-.007	.837	.019	.028	.893	.460	.770	.930	.235	.447	.948	.183	.383	.948	.195	.408
4	.949	.102	.215	.795	.279	.397	.929	.269	.507	.930	.458	.869	.914	.161	.288	.914	.331	.591
5	.627	.375	.479	.614	.335	.426	.527	.298	.374	.895	.397	.669	.810	.491	.710	.931	.535	1.022
6	.845	-.118	-.180	.698	.102	.135	.842	.001	.002	.860	.305	.475	.931	-.097	-.186	.862	.164	.257
7	.729	.022	.029	.356	.351	.451	.772	-.092	-.128	.754	.233	.319	.672	.523	.680	.776	.341	.475
8	.508	.356	.446	.822	.210	.309	.268	.173	.232	.860	.410	.640	.862	.365	.572	.948	.405	.846
9	.864	.243	.382	.587	.370	.468	.754	.455	.622	.930	.262	.498	.845	.233	.354	.828	.313	.463
10	.847	-.047	-.072	.182	.261	.382	.667	-.054	-.070	.714	.259	.344	.741	.269	.364	.707	.157	.208
11	.421	.266	.335	.600	.337	.427	.352	.302	.389	.607	.381	.484	.474	.280	.351	.649	.139	.178
12	.857	.176	.272	.851	.264	.405	.944	.292	.597	.804	.509	.731	.684	.608	.795	.893	.555	.931
13	.691	.320	.420	.565	.325	.409	.558	.218	.274	.582	.426	.538	.611	.591	.752	.782	.351	.492
14	.529	.608	.763	.523	.283	.355	.212	.421	.594	.804	.289	.415	.891	.536	.893	.852	.424	.651
15	.619	.318	.405	.561	.451	.568	.720	.325	.434	.808	.386	.557	.740	.425	.575	.800	.428	.611
Session 2																		
1	.982	.142	.429	.955	.168	.365	1.00	0.000	0.000	.965	-.068	-.161	.964	.143	.338	.947	.388	.807
2	.860	.175	.273	.929	.251	.474	.927	.287	.539	.860	.378	.588	.982	.251	.757	.912	.092	.163
3	.754	.244	.333	.786	.508	.714	.818	.305	.446	.877	.216	.349	.930	.197	.374	.895	.061	.103
4	.561	.499	.628	.452	.362	.455	.636	.024	.030	.877	.561	.906	.912	.545	.968	.895	.109	.184
5	.911	.078	.139	.882	.165	.269	.745	.006	.008	.860	.475	.740	.842	.240	.362	.754	.303	.415
6	.947	.183	.381	.907	.054	.094	.964	.025	.058	.912	.177	.314	.860	.189	.295	.877	.058	.094
7	.825	.233	.344	.800	.281	.402	.818	.367	.536	.544	.397	.499	.649	.229	.295	.789	.098	.139
8	.661	.342	.442	.765	.056	.078	.818	.158	.230	.754	.448	.613	.825	.381	.561	.912	.389	.690
9	.772	.170	.236	.638	.191	.245	.691	.140	.183	.789	.308	.435	.684	.295	.386	.737	.089	.121
10	.912	.139	.247	.938	.208	.410	.982	.130	.388	.754	.399	.546	.807	.264	.381	.754	.210	.288
11	.807	.150	.216	.688	.069	.090	.873	.053	.085	.643	.469	.602	.526	.288	.361	.500	.262	.328
12	.895	.141	.238	.978	.125	.352	.909	.254	.447	.842	.291	.440	.789	.661	.934	.750	.409	.558
13	.842	.204	.309	.756	.168	.230	.909	.293	.515	.782	.345	.483	.750	.646	.881	.679	.512	.668
14	.912	.230	.408	.911	.240	.425	.891	.303	.506	.800	.549	.784	.836	.508	.761	.782	.342	.479
15	.673	.474	.616	.510	.172	.216	.709	.210	.278	.792	.529	.750	.635	.411	.526	.741	.466	.631

Note. For Session 1, computer data are from Group 1 and paper-and-pencil data are from Group 2; for Session 2, computer data are from Group 2 and paper-and-pencil data are from Group 1.

[a]For Session 1 paper-and-pencil data, the mode designation refers to which computer administration mode followed the paper-and-pencil administration (Group 2); for Session 2, this designation indicates that the paper-and-pencil data were calculated from tests that followed the particular computer administration mode listed (Group 1).

Table B-2. Item Difficulties (p), Point-Biserial Correlations (rpb), and Biserial Correlations (rbis), for Version B of the PC Subtest by Session and Test Administration Medium

Session and item	Computer									Paper-and-Pencil								
	Mode 1			Mode 2			Mode 3			Mode 1[a]			Mode 2[a]			Mode 3[a]		
	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis
Session 1																		
1	.946	.082	.170	.940	.371	.740	.945	.004	.008	.964	.191	.451	.982	.050	.149	1.00	0.000	0.000
2	.911	.007	.013	.816	.289	.421	.929	.243	.460	.929	.222	.420	.909	.315	.553	.931	-.018	-.035
3	.714	.117	.156	.673	.242	.315	.589	-.058	-.074	.857	.401	.622	.909	-.076	-.133	.914	.237	.424
4	.446	.157	.197	.510	-.006	-.008	.446	.135	.170	.804	.242	.347	.636	.343	.439	.828	.030	.044
5	.804	.069	.099	.786	.067	.095	.821	.190	.278	.893	.397	.665	.800	-.108	-.154	.741	.088	.119
6	.857	.344	.534	.768	.157	.216	.875	.350	.562	.911	.241	.426	.927	.178	.335	.966	.008	.020
7	.750	-.047	-.065	.647	.058	.074	.696	.141	.186	.768	-.013	-.019	.836	.035	.052	.793	.263	.372
8	.315	.210	.275	.537	.129	.161	.268	.232	.311	.714	.342	.454	.782	.258	.362	.879	.191	.310
9	.556	-.073	-.092	.423	-.257	-.324	.696	.177	.233	.696	.024	.032	.618	.289	.368	.690	.133	.174
10	.481	.120	.151	.725	.006	.009	.500	.456	.571	.893	.261	.438	.909	.190	.334	.983	-.029	-.089
11	.865	.226	.356	.837	.178	.267	.846	-.146	-.223	.800	.219	.313	.731	.167	.224	.849	-.009	-.014
12	.755	.160	.219	.868	.260	.412	.660	.156	.201	.855	.200	.309	.885	.422	.693	.907	.137	.239
13	.638	.242	.310	.917	.361	.650	.692	.195	.256	.830	.305	.453	.780	.026	.037	.925	.212	.393
14	.844	.199	.302	.875	.596	.958	.878	.317	.513	.904	.393	.680	.872	.372	.594	.846	.293	.446
15	.553	.516	.649	.522	.467	.586	.500	.072	.090	.681	.434	.566	.614	.333	.424	.700	.099	.131
Session 2																		
1	.982	.118	.357	1.00	0.000	0.000	.964	.466	1.100	.948	.163	.341	.948	.215	.449	.982	.162	.488
2	.887	.161	.265	.783	.006	.008	.929	.158	.299	.914	.384	.686	.759	.222	.304	.893	.076	.128
3	.929	.505	.953	.958	.185	.415	.982	.103	.309	.879	-.072	-.117	.862	-.009	-.014	.893	.198	.332
4	.946	.389	.804	.958	.309	.693	.911	.473	.835	.741	.392	.531	.586	.294	.372	.696	.340	.447
5	.821	.411	.603	.826	.545	.804	.893	.206	.345	.810	.164	.238	.776	-.056	-.078	.857	.272	.422
6	.929	.226	.427	.945	-.064	-.132	.964	-.151	-.358	.879	.246	.399	.897	.063	.107	.875	.529	.850
7	.714	.284	.377	.667	.211	.273	.732	.316	.425	.719	.165	.220	.724	.283	.379	.696	.233	.306
8	.727	.201	.269	.833	.218	.325	.821	.396	.582	.776	.234	.327	.690	.178	.233	.732	.100	.135
9	.946	.319	.659	.852	.388	.596	.909	.086	.151	.544	.294	.369	.517	.379	.475	.709	.271	.359
10	.839	-.052	-.078	.667	.402	.521	.768	.114	.158	.912	-.030	-.053	.914	.196	.349	.909	.308	.540
11	.500	.537	.673	.566	.317	.400	.589	.321	.407	.736	.348	.469	.807	.040	.057	.745	.006	.008
12	.786	.364	.512	.849	.414	.633	.893	.614	1.029	.811	.338	.489	.839	.226	.340	.836	.199	.299
13	.643	.199	.256	.642	.355	.456	.786	.397	.559	.765	-.013	-.018	.768	.060	.083	.745	.362	.491
14	.778	.505	.705	.849	.510	.779	.745	.515	.700	.755	.487	.667	.786	.362	.510	.833	.448	.668
15	.768	.286	.395	.717	.252	.335	.821	.348	.511	.673	.348	.452	.585	.270	.342	.660	.208	.269

Note. For Session 1, computer data are from Group 1 and paper-and-pencil data are from Group 2; for Session 2, computer data are from Group 2 and paper-and-pencil data are from Group 1.

[a]For Session 1 paper-and-pencil data, the mode designation refers to which computer administration mode followed the paper-and-pencil administration (Group 2); for Session 2 this designation indicates that the paper-and-pencil data were calculated from tests that followed the particular computer administration mode listed (Group 1).

Table B-3. Item Difficulties (p), Point-Biserial Correlations (rpb), and Biserial Correlations (rbis), Within Session 1 for Version A of the AS Subtest by Form and Test Administration Medium

	Computer									Paper-and-Pencil								
	Form 11			Form 12			Form 13			Form 11			Form 12			Form 13		
Item	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis
1	.864	.560	.881	.907	.392	.685	.930	.366	.695	.860	.494	.770	.931	.449	.858	.931	.051	.098
2	.763	.225	.310	.963	.212	.495	.982	.167	.506	.912	.239	.424	.793	.134	.190	.914	.205	.365
3	.810	.316	.457	.815	-.110	-.161	.789	.398	.562	.860	.190	.296	.914	.325	.580	.845	.192	.292
4	.847	.330	.503	.963	.147	.342	.912	.204	.363	.860	.317	.494	.776	.397	.554	.845	.336	.510
5	.763	.270	.372	.796	.488	.694	.877	.173	.279	.754	.229	.313	.828	.472	.698	.862	.172	.270
6	.864	.198	.311	.778	.058	.081	.930	.280	.531	.842	.108	.163	.672	.324	.422	.879	.261	.423
7	.593	.192	.243	.833	.380	.567	.772	.065	.090	.649	.382	.492	.672	.383	.497	.793	.320	.453
8	.746	.598	.813	.685	.503	.658	.877	.375	.606	.930	.278	.528	.741	.210	.284	.879	.372	.603
9	.847	.351	.536	.667	.496	.643	.842	.275	.415	.877	.066	.107	.702	.563	.743	.776	.202	.281
10	.729	.373	.501	.815	.361	.525	.649	.591	.761	.719	.273	.365	.776	.407	.567	.655	.462	.596
11	.814	.351	.510	.870	.482	.767	.684	.550	.719	.719	.166	.221	.719	.606	.808	.707	.266	.352
12	.695	.345	.454	.679	.145	.189	.649	.233	.300	.684	.445	.581	.776	.259	.361	.621	.165	.210
13	.576	.511	.645	.741	.463	.627	.333	.337	.436	.750	.429	.585	.759	.623	.855	.672	.496	.644
14	.763	.083	.114	.685	.164	.214	.807	.445	.641	.772	.051	.071	.672	.178	.231	.672	.514	.668
15	.712	.372	.493	.463	.339	.425	.702	.483	.638	.786	.421	.592	.707	.490	.649	.649	.618	.796
16	.603	.321	.407	.630	.392	.501	.526	.213	.267	.667	.417	.541	.684	.377	.493	.655	.193	.249
17	.695	.227	.298	.679	.434	.566	.632	.476	.609	.702	.194	.255	.690	.267	.349	.534	.463	.581
18	.644	.210	.270	.611	.264	.335	.719	.417	.556	.684	.336	.440	.603	.427	.541	.638	.326	.418
19	.576	.420	.530	.815	.451	.657	.554	.635	.799	.536	.664	.834	.724	.631	.844	.724	.373	.499
20	.448	.092	.116	.500	.322	.404	.607	.64[illegible]	.818	.607	.529	.672	.603	.173	.220	.448	.538	.676
21	.542	.673	.846	.204	-.006	-.009	.607	.209	.265	.518	.686	.860	.086	-.024	-.043	.737	.310	.418
22	.759	.173	.238	.500	.520	.651	.393	.336	.427	.789	.164	.232	.431	.336	.424	.482	.434	.544
23	.632	.354	.453	.611	.371	.472	.518	.421	.528	.625	.338	.432	.672	.170	.221	.439	.349	.439
24	.368	.374	.478	.660	.400	.518	.436	.479	.603	.327	.531	.690	.439	.436	.549	.561	.282	.355
25	.456	.224	.282	.528	.659	.827	.473	.577	.724	.625	.302	.386	.544	.231	.290	.421	.138	.175

Note. Computer data are from Group 1 and paper-and-pencil data are from Group 2.

Table B-4. Item Difficulties (p), Point-Biserial Correlations (rpb), and Biserial Correlations (rbis), Within Session 1 for Version B of the AS Subtest by Form and Test Administration Medium

	Computer									Paper-and-Pencil								
	Form 11			Form 12			Form 13			Form 11			Form 12			Form 13		
Item	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis
1	.875	.217	.349	.786	.430	.604	.889	.328	.545	.800	.296	.424	.818	.373	.545	.966	.144	.345
2	.768	.260	.360	.839	.565	.850	.964	.266	.628	.855	.449	.693	.891	.231	.385	.966	.235	.561
3	.893	.258	.433	.873	.486	.777	.839	.496	.747	.891	.356	.594	.909	.416	.731	.983	.210	.641
4	.804	.502	.721	.839	.486	.731	.875	.432	.694	.815	.423	.615	.909	.173	.305	.897	.186	.315
5	.625	.471	.601	.857	.256	.397	.839	.170	.255	.782	.360	.505	.727	.396	.530	.879	.122	.198
6	.964	.266	.629	.750	.419	.571	.836	.301	.451	.964	.252	.591	.727	.478	.641	.931	.130	.248
7	.821	.315	.463	.714	.463	.615	.804	.290	.417	.891	.419	.698	.745	.294	.399	.828	.317	.469
8	.818	.117	.170	.696	.224	.294	.839	.257	.387	.727	.426	.571	.691	.463	.607	.879	.529	.858
9	.714	.609	.810	.727	.714	.957	.696	.479	.630	.764	.201	.277	.764	.385	.531	.810	.165	.239
10	.836	.645	.966	.857	.571	.886	.679	.461	.600	.800	.535	.764	.727	.351	.470	.724	.335	.448
11	.786	.458	.644	.625	.601	.768	.732	.327	.439	.873	.386	.617	.745	.535	.727	.741	.420	.569
12	.764	.316	.435	.696	.489	.644	.679	.301	.392	.745	.447	.607	.709	.292	.387	.741	.420	.569
13	.745	.215	.291	.518	.558	.699	.679	.592	.771	.818	.537	.785	.764	.548	.756	.793	.515	.730
14	.709	.041	.054	.679	.171	.223	.661	.264	.341	.745	.274	.372	.764	.254	.350	.759	.285	.390
15	.696	.511	.673	.800	.401	.574	.536	.301	.378	.673	.208	.271	.709	.443	.587	.707	.312	.413
16	.732	.241	.325	.589	.222	.281	.509	.221	.277	.727	.293	.392	.745	.432	.586	.649	.242	.311
17	.589	.664	.840	.821	.571	.838	.685	.510	.667	.545	.631	.792	.800	.552	.788	.638	.165	.212
18	.571	.406	.512	.607	.539	.684	.745	.422	.572	.727	.214	.287	.673	.537	.698	.719	.399	.533
19	.607	.371	.472	.714	.351	.466	.643	.524	.673	.727	.508	.681	.782	.233	.327	.690	.513	.672
20	.400	.147	.186	.589	.280	.354	.400	.565	.716	.527	.334	.419	.655	.419	.541	.517	.576	.722
21	.727	.309	.414	.655	.396	.511	.655	.445	.574	.673	.419	.544	.636	.056	.072	.638	.391	.502
22	.818	.207	.302	.554	.431	.542	.463	.377	.473	.778	.335	.468	.545	.602	.757	.500	.239	.300
23	.545	.589	.739	.143	.128	.198	.611	.349	.444	.537	.497	.624	.093	-.009	-.016	.776	.208	.290
24	.382	.288	.367	.429	.399	.502	.455	.637	.800	.426	.321	.405	.585	.346	.438	.517	.492	.617
25	.564	.349	.440	.446	.282	.354	.519	.404	.507	.519	.532	.668	.472	.399	.501	.466	.467	.586

Note. Computer data are from Group 1 and paper-and-pencil data are from Group 2.

Table B-5. Item Difficulties (p), Point-Biserial Correlations (rpb), and Biserial Correlations (rbis), Within Session 2 for Version A of the AS Subtest by Form and Test Administration Medium

	Computer									Paper-and-Pencil								
	Form 11			Form 12			Form 13			Form 11			Form 12			Form 13		
Item	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis
1	.789	.321	.453	.754	.565	.772	.944	.367	.749	.877	.373	.602	.895	.492	.830	.877	.380	.614
2	.842	.572	.865	.877	.385	.622	.964	.186	.436	.772	.349	.485	.895	.443	.746	.930	.309	.586
3	.912	.332	.590	.929	.424	.800	.891	-.045	-.075	.842	.346	.523	.860	.583	.908	.842	.464	.701
4	.825	.256	.378	.842	.479	.724	.889	.260	.432	.825	.264	.389	.860	.459	.716	.912	.254	.452
5	.789	.422	.596	.772	.447	.621	.873	.096	.154	.842	.449	.678	.825	.464	.683	.737	.400	.540
6	.982	-.033	-.100	.684	.319	.418	.891	.390	.651	.965	.086	.205	.772	.455	.632	.860	.457	.712
7	.875	.403	.648	.789	.244	.344	.818	.191	.279	.649	.558	.718	.772	.227	.315	.842	.216	.326
8	.737	.219	.295	.732	.258	.347	.891	.522	.871	.825	.264	.389	.737	.205	.277	.842	.501	.758
9	.789	.381	.538	.737	.497	.671	.855	.101	.155	.789	.395	.558	.719	.648	.864	.754	.392	.535
10	.772	.492	.682	.732	.421	.566	.818	.248	.363	.789	.602	.850	.912	.615	1.093	.772	.384	.533
11	.912	.318	.565	.719	.505	.673	.800	.261	.373	.649	.666	.858	.737	.704	.950	.754	.376	.514
12	.772	.284	.394	.684	.387	.505	.764	.362	.499	.754	.377	.516	.737	.343	.463	.684	.416	.544
13	.807	.382	.550	.737	.407	.550	.800	.432	.617	.754	.368	.504	.825	.363	.534	.719	.372	.496
14	.719	.280	.373	.714	.396	.526	.727	.336	.450	.632	.170	.218	.702	.245	.324	.719	.281	.375
15	.667	.310	.402	.772	.391	.543	.673	.264	.343	.772	.261	.362	.789	.409	.577	.684	.605	.791
16	.807	.248	.357	.702	.341	.450	.618	.077	.099	.702	.445	.587	.614	.236	.300	.554	.292	.368
17	.614	.458	.583	.842	.299	.452	.600	.403	.512	.649	.371	.478	.737	.282	.381	.579	.627	.792
18	.737	.368	.496	.607	.566	.719	.709	.391	.519	.579	.367	.463	.632	.499	.639	.719	.612	.816
19	.702	.488	.643	.772	.299	.416	.709	.580	.769	.579	.579	.731	.825	.443	.653	.754	.609	.833
20	.596	.254	.322	.737	.155	.209	.509	.510	.640	.526	.461	.579	.649	.239	.308	.446	.533	.670
21	.702	.216	.284	.649	.193	.248	.648	.296	.381	.544	.434	.545	.123	.313	.506	.632	.382	.489
22	.750	.272	.371	.526	.545	.684	.537	.253	.318	.719	.299	.398	.544	.438	.551	.526	.485	.608
23	.579	.443	.559	.105	.143	.241	.722	.392	.523	.719	.433	.578	.737	.205	.277	.702	.405	.535
24	.491	.413	.518	.554	.147	.185	.519	.559	.701	.446	.199	.251	.491	.480	.602	.456	.644	.809
25	.579	.434	.548	.564	.139	.175	.426	.418	.527	.526	.263	.330	.456	.215	.270	.474	.557	.699

Note. Computer data are from Group 2 and paper-and-pencil data are from Group 1.

Table B-6. Item Difficulties (p), Point-Biserial Correlations (rpb), and Biserial Correlations (rbis), Within Session 2 for Version B of the AS Subtest by Form and Test Administration Medium

	Computer									Paper-and-Pencil								
	Form 11			Form 12			Form 13			Form 11			Form 12			Form 13		
Item	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis
1	.789	.321	.453	.754	.565	.772	.944	.367	.749	.776	.513	.715	.828	.055	.081	.764	.263	.362
2	.842	.572	.865	.877	.385	.622	.964	.186	.436	.828	.213	.315	.948	.162	.339	.945	.189	.388
3	.912	.332	.590	.929	.424	.800	.891	-.045	-.075	.862	.461	.721	.931	.472	.901	.891	.160	.266
4	.825	.256	.378	.842	.479	.724	.889	.260	.432	.862	.269	.421	.931	.252	.482	.945	.078	.161
5	.789	.422	.596	.772	.447	.621	.873	.096	.154	.638	.422	.541	.810	.392	.566	.855	.105	.161
6	.982	-.033	-.100	.684	.319	.418	.891	.390	.651	.931	.081	.155	.707	.190	.251	.891	.366	.611
7	.875	.403	.648	.789	.244	.344	.818	.191	.279	.793	.389	.552	.879	.218	.354	.891	.310	.518
8	.737	.219	.295	.732	.258	.347	.891	.522	.871	.776	.532	.742	.754	.403	.551	.873	.277	.443
9	.789	.381	.538	.737	.497	.671	.855	.101	.155	.828	.366	.542	.842	.532	.804	.800	.287	.410
10	.772	.492	.682	.732	.421	.566	.818	.248	.363	.724	.337	.451	.828	.412	.610	.727	.364	.488
11	.912	.318	.565	.719	.505	.673	.800	.261	.373	.897	.234	.396	.655	.310	.400	.855	.262	.404
12	.772	.284	.394	.684	.387	.505	.764	.362	.499	.810	.273	.395	.707	.060	.080	.636	.317	.406
13	.807	.382	.550	.737	.407	.550	.800	.432	.617	.754	.394	.538	.741	.370	.501	.582	.451	.569
14	.719	.280	.373	.714	.396	.526	.727	.336	.450	.655	.088	.114	.638	.487	.624	.836	.320	.480
15	.667	.310	.402	.772	.391	.543	.673	.264	.343	.672	.497	.646	.776	.467	.651	.491	.302	.378
16	.807	.248	.357	.702	.341	.450	.618	.077	.099	.603	.487	.618	.655	.388	.501	.564	.335	.421
17	.614	.458	.583	.842	.299	.452	.600	.403	.512	.638	.356	.456	.741	.427	.578	.618	.382	.487
18	.737	.368	.496	.607	.566	.719	.709	.391	.519	.649	.180	.231	.655	.353	.456	.745	.279	.378
19	.702	.488	.643	.772	.299	.416	.709	.580	.769	.638	.482	.617	.741	.475	.643	.636	.506	.648
20	.596	.254	.322	.737	.155	.209	.509	.510	.640	.500	.220	.275	.552	.373	.468	.600	.641	.813
21	.702	.216	.284	.649	.193	.248	.648	.296	.381	.649	.238	.307	.793	.195	.276	.545	.485	.610
22	.750	.272	.371	.526	.545	.684	.537	.253	.318	.719	.223	.298	.638	.514	.659	.455	.344	.432
23	.579	.443	.559	.105	.143	.241	.722	.392	.523	.596	.453	.574	.190	.212	.306	.582	.136	.172
24	.491	.413	.518	.554	.147	.185	.519	.559	.701	.439	.355	.447	.571	.310	.391	.436	.389	.490
25	.579	.434	.548	.564	.139	.175	.426	.418	.527	.404	.324	.410	.614	.546	.695	.545	.366	.459

Note. Computer data are from Group 2 and paper-and-pencil data are from Group 1.

Table B-7. Item Difficulties (p), Point-Biserial Correlations (rpb), and Biserial Correlations (rbis), Within Session 1 for Version A of the MC Subtest by Form and Test Administration Medium

	Computer									Paper-and-Pencil								
	Form 11			Form 12			Form 13			Form 11			Form 12			Form 13		
Item	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis
1	.898	.345	.587	1.00	0.000	0.000	.912	.456	.809	.930	.137	.260	.931	.376	.718	.879	.377	.612
2	.831	.101	.150	.815	.265	.386	.930	.260	.493	.895	.230	.388	.862	.343	.536	.879	.044	.072
3	.898	.381	.649	.741	.265	.359	.842	.366	.553	.912	-.072	-.128	.897	.410	.694	.862	-.006	-.009
4	.881	.121	.197	.830	.032	.048	.732	.419	.564	.895	.074	.125	.914	.121	.217	.793	.050	.070
5	.729	.242	.325	.906	.247	.429	.877	.392	.632	.877	.059	.095	.862	.418	.654	.914	-.000	-.001
6	.508	.296	.371	.365	.124	.159	.772	.346	.480	.579	.067	.084	.776	.110	.153	.845	.069	.104
7	.627	.404	.516	.852	.453	.696	.875	.329	.528	.842	.411	.621	.741	.500	.677	.810	.346	.501
8	.763	.303	.417	.642	.262	.336	.684	.417	.545	.877	.205	.330	.552	.449	.565	.724	.396	.530
9	.621	.422	.537	.519	.335	.420	.772	.162	.225	.684	.464	.607	.509	.615	.771	.845	.122	.185
10	.644	.477	.613	.833	.411	.613	.930	.312	.592	.860	.368	.574	.810	.363	.525	.893	.439	.736
11	.644	.342	.440	.556	.491	.618	.702	.238	.314	.807	.066	.095	.776	.170	.237	.810	.373	.539
12	.569	.465	.586	.585	.492	.622	.789	.309	.436	.825	.215	.317	.690	.560	.733	.724	-.028	-.037
13	.695	.277	.364	.519	.140	.176	.714	.183	.243	.667	.236	.306	.724	.265	.355	.741	.253	.342
14	.712	.453	.602	.593	.289	.366	.807	.202	.292	.719	.384	.512	.690	.418	.547	.810	.256	.370
15	.661	.432	.558	.722	.188	.252	.684	.316	.414	.754	.423	.578	.707	.280	.371	.828	.164	.243
16	.610	.311	.395	.611	.455	.578	.456	.388	.487	.772	.330	.458	.707	.371	.490	.776	.210	.292
17	.661	.157	.204	.630	.354	.453	.464	.124	.156	.614	.181	.230	.397	.092	.117	.466	.407	.511
18	.627	.212	.270	.593	.439	.556	.491	.204	.256	.649	.446	.574	.603	.417	.529	.603	.308	.391
19	.525	.513	.643	.630	.345	.441	.596	.209	.265	.702	.196	.259	.724	.340	.455	.632	.186	.238
20	.552	.350	.440	.630	.612	.783	.579	.269	.339	.702	.069	.091	.534	.579	.727	.561	.167	.210
21	.305	.350	.459	.574	.371	.468	.719	.109	.146	.554	.374	.470	.603	.370	.469	.518	.302	.379
22	.456	.316	.397	.547	.388	.488	.250	.195	.266	.625	.409	.522	.552	.304	.382	.375	.225	.287
23	.517	.452	.566	.358	-.044	-.056	.393	.325	.413	.518	.268	.336	.439	.481	.605	.418	.318	.401
24	.400	.370	.469	.481	.419	.525	.421	.388	.490	.259	.245	.331	.456	.446	.561	.444	.435	.547
25	.382	.253	.322	.396	.418	.530	.404	.247	.313	.593	.244	.309	.421	.278	.351	.315	.228	.298

Note. Computer data are from Group 1 and paper-and-pencil data are from Group 2.

Table B-8. Item Difficulties (p), Point-Biserial Correlations (rpb), and Biserial Correlations (rbis), Within Session 1 for Version B of the MC Subtest by Form and Test Administration Medium

	Computer									Paper-and-Pencil								
	Form 11			Form 12			Form 13			Form 11			Form 12			Form 13		
Item	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis
1	.836	.602	.902	.764	.225	.310	.786	.322	.453	.964	.062	.148	.709	.337	.447	.793	.229	.324
2	.964	.054	.126	.839	.167	.251	.821	.264	.387	.929	.112	.212	.782	.348	.488	.966	.109	.260
3	.909	.431	.758	.946	.272	.562	.893	.153	.256	.964	.010	.024	1.00	0.00	0.00	.931	.240	.458
4	.836	.305	.457	.839	.321	.483	.625	.569	.726	.929	.306	.577	.800	.295	.422	.862	.307	.481
5	.764	.464	.639	.857	.032	.050	.804	.250	.358	.821	.043	.063	.836	.235	.352	.862	.016	.024
6	.481	.136	.171	.455	.021	.026	.696	.125	.164	.571	.091	.115	.818	.230	.335	.897	.123	.208
7	.839	.358	.539	.907	.502	.877	.732	.300	.404	.875	.164	.263	.909	.224	.393	.845	.276	.419
8	.768	.477	.660	.732	.279	.375	.607	.094	.120	.929	.208	.393	.618	.308	.393	.793	.128	.182
9	.768	.424	.586	.714	.313	.416	.696	.285	.375	.786	.338	.476	.745	.190	.257	.862	.159	.249
10	.691	.385	.506	.818	.213	.312	.786	.322	.453	.750	.287	.392	.818	.302	.441	.845	.421	.639
11	.643	.392	.504	.607	.475	.603	.821	.447	.656	.714	.327	.434	.673	.149	.194	.897	.156	.264
12	.600	.398	.505	.509	.371	.465	.750	.207	.282	.857	.516	.800	.618	.260	.331	.793	.067	.094
13	.655	.360	.464	.774	.119	.165	.564	.171	.215	.750	.372	.508	.727	.317	.424	.672	.312	.406
14	.643	.448	.575	.536	.320	.401	.691	.142	.186	.714	.067	.089	.582	.388	.490	.707	.238	.315
15	.589	.287	.363	.625	.081	.103	.964	.038	.091	.714	.189	.251	.618	.042	.054	.724	.184	.246
16	.655	.211	.272	.661	.301	.389	.556	.141	.177	.732	.254	.341	.600	.095	.120	.448	.305	.384
17	.589	.385	.487	.661	.310	.401	.673	.165	.214	.607	.549	.698	.709	.285	.378	.655	.178	.230
18	.714	.262	.348	.774	.463	.643	.696	.253	.332	.764	.091	.126	.600	.353	.447	.586	.121	.153
19	.509	.481	.602	.625	.545	.696	.500	.100	.125	.691	.155	.203	.691	.348	.456	.500	.297	.372
20	.582	.387	.489	.661	.310	.401	.589	.223	.282	.709	.300	.397	.722	.447	.597	.552	.262	.330
21	.519	.332	.416	.352	.307	.395	.357	.114	.146	.491	.338	.423	.481	-.162	-.203	.534	.287	.361
22	.519	.376	.472	.611	.145	.185	.321	.180	.235	.673	.288	.374	.585	.315	.399	.466	.090	.113
23	.396	.338	.429	.537	.387	.486	.518	.377	.472	.556	.467	.587	.642	.323	.414	.517	.239	.299
24	.352	.156	.201	.528	.397	.498	.509	.181	.227	.226	.291	.404	.491	.406	.509	.431	.324	.408
25	.463	.291	.365	.377	.434	.554	.345	.262	.338	.566	.126	.159	.320	.459	.598	.362	.422	.541

Note. Computer data are from Group 1 and paper-and-pencil data are from Group 2.

Table B-9. Item Difficulties (p), Point-Biserial Correlations (rpb), and Biserial Correlations (rbis), Within Session 2 for Version A of the MC Subtest by Form and Test Administration Medium

	Computer									Paper-and-Pencil								
	Form 11			Form 12			Form 13			Form 11			Form 12			Form 13		
Item	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis
1	.965	.086	.204	.684	.327	.427	.764	.189	.260	.912	.626	1.111	.982	.222	.672	.965	.301	.715
2	.912	.169	.300	.768	.383	.530	.945	.052	.107	.877	.360	.582	.860	.281	.437	.877	.190	.307
3	.912	.251	.446	.982	.059	.177	.927	.342	.642	.912	.524	.931	.807	.247	.355	.842	.228	.344
4	.912	.218	.388	.842	.280	.423	.855	.355	.549	.825	.259	.382	.877	.137	.221	.649	.378	.486
5	.789	.170	.240	.804	.341	.490	.852	-.055	-.084	.895	.460	.776	.912	.228	.405	.719	.357	.475
6	.491	.242	.303	.804	.209	.300	.800	.236	.337	.474	-.236	-.296	.632	.363	.464	.737	.129	.174
7	.877	.215	.346	.860	.411	.640	.836	.209	.314	.807	.366	.528	.860	.157	.244	.912	.270	.480
8	.877	.158	.255	.544	.282	.354	.782	.128	.180	.877	.324	.523	.684	.195	.255	.667	.148	.192
9	.768	.336	.464	.679	.271	.353	.891	.098	.164	.702	.337	.444	.719	.367	.490	.825	.357	.526
10	.807	.474	.683	.737	.419	.566	.891	.333	.555	.649	.376	.483	.825	.275	.405	.930	.159	.301
11	.649	.298	.383	.661	.317	.409	.855	.174	.269	.754	.456	.623	.772	.220	.305	.877	.324	.523
12	.754	.366	.500	.679	.472	.616	.836	.167	.250	.719	.420	.559	.737	.462	.624	.807	.093	.134
13	.719	.422	.562	.667	.339	.440	.764	.303	.417	.649	.197	.253	.807	.203	.292	.930	.066	.126
14	.719	.152	.203	.643	.306	.393	.655	.269	.347	.754	.456	.623	.702	.507	.668	.702	.059	.078
15	.719	.290	.387	.554	.290	.365	.836	.139	.208	.632	.268	.343	.754	.169	.231	.737	.423	.571
16	.702	.379	.501	.691	.064	.084	.491	.237	.297	.702	.373	.492	.667	.377	.489	.561	.267	.336
17	.684	.447	.584	.768	.276	.381	.636	.226	.290	.561	.380	.478	.702	.288	.380	.544	.159	.199
18	.807	.074	.107	.561	.249	.314	.618	.229	.292	.667	.201	.261	.754	.383	.524	.649	.303	.390
19	.719	.163	.217	.554	.442	.556	.519	.237	.297	.596	.352	.446	.737	.239	.322	.596	.318	.403
20	.719	.301	.401	.772	.395	.548	.618	.342	.436	.667	.444	.576	.702	.406	.536	.649	.388	.500
21	.526	.376	.472	.386	.011	.014	.545	.325	.409	.368	.472	.604	.526	.341	.428	.579	.156	.197
22	.719	.205	.273	.582	.287	.362	.500	.089	.112	.579	.283	.357	.596	.427	.541	.316	.311	.406
23	.632	.417	.534	.589	.318	.402	.519	.252	.316	.544	.230	.289	.404	.341	.432	.386	.323	.411
24	.298	.156	.206	.446	.420	.528	.473	.290	.363	.351	.331	.427	.544	.325	.408	.526	.240	.301
25	.554	.195	.245	.333	.417	.540	.370	.401	.512	.579	.433	.547	.421	.320	.403	.351	.299	.385

Note. Computer data are from Group 2 and paper-and-pencil data are from Group 1.

Table B-10. Item Difficulties (p), Point-Biserial Correlations (rpb), and Biserial Correlations (rbis), Within Session 2 for Version B of the MC Subtest by Form and Test Administration Medium

	Computer									Paper-and-Pencil								
	Form 11			Form 12			Form 13			Form 11			Form 12			Form 13		
Item	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis
1	.911	.290	.513	.914	.110	.196	.929	.261	.494	.948	.273	.572	.789	.224	.316	.836	.348	.522
2	.857	.253	.392	.842	.300	.454	.929	.030	.056	.845	.148	.225	.842	.236	.356	.909	.112	.197
3	.911	.221	.391	.879	.340	.551	.804	.179	.257	.897	.246	.416	1.00	0.000	0.000	.891	.438	.730
4	.929	-.003	-.005	.914	.229	.409	.839	.182	.273	.879	.266	.432	.825	.179	.263	.745	.259	.352
5	.873	.234	.374	.897	.343	.581	.946	-.029	-.060	.707	.346	.458	.895	.078	.132	.891	.182	.304
6	.536	.213	.268	.810	.046	.066	.732	-.106	-.143	.517	.091	.114	.719	.227	.303	.818	.319	.466
7	.818	.319	.466	.754	.370	.506	.855	.324	.500	.776	.150	.210	.947	.354	.736	.891	.182	.304
8	.873	.201	.322	.621	.253	.322	.696	.345	.454	.845	.238	.362	.649	.223	.287	.709	.112	.149
9	.679	.299	.390	.586	.384	.486	.839	.398	.599	.793	.224	.318	.719	.313	.417	.764	.190	.261
10	.804	.420	.603	.845	.129	.196	.946	.234	.483	.672	.565	.734	.825	.402	.592	.927	.288	.542
11	.800	.099	.142	.862	.142	.222	.732	.408	.549	.638	.366	.470	.772	.176	.244	.873	.168	.268
12	.786	.204	.287	.741	.474	.641	.750	-.049	-.067	.655	.199	.256	.614	.381	.485	.800	.254	.363
13	.661	.282	.365	.696	.306	.403	.804	.351	.503	.586	.334	.423	.772	.452	.628	.800	.201	.287
14	.714	.225	.299	.690	.403	.528	.804	.110	.157	.690	.325	.426	.579	.406	.513	.782	.131	.184
15	.804	.368	.528	.655	.495	.639	.833	.146	.218	.741	.156	.211	.561	.054	.068	.764	.290	.400
16	.786	.314	.442	.655	.222	.287	.750	.335	.456	.690	.325	.426	.632	.393	.502	.545	.340	.427
17	.643	.172	.221	.397	.166	.211	.482	.381	.478	.603	.394	.500	.667	.353	.458	.618	.173	.220
18	.679	.255	.333	.655	.343	.442	.589	.153	.194	.672	.094	.122	.596	.471	.597	.618	.261	.333
19	.673	.404	.525	.789	.263	.371	.607	.105	.133	.638	.200	.256	.737	.349	.471	.564	.045	.057
20	.714	-.064	-.085	.621	.456	.581	.571	.123	.155	.586	.206	.261	.684	.609	.796	.491	.390	.489
21	.643	.256	.328	.614	.279	.356	.571	.225	.284	.552	.315	.396	.351	.355	.457	.436	.194	.244
22	.732	.394	.530	.569	.050	.063	.364	.092	.118	.483	.546	.684	.632	.329	.421	.327	.272	.354
23	.491	.206	.258	.414	.348	.440	.446	.255	.320	.431	.535	.674	.632	.320	.409	.691	.034	.045
24	.291	.193	.256	.439	.462	.582	.464	.305	.383	.431	.384	.484	.561	.214	.269	.491	.268	.336
25	.625	.456	.583	.431	.201	.253	.304	.192	.252	.404	.419	.531	.386	.454	.578	.400	.403	.511

Note. Computer data are from Group 2 and paper-and-pencil data are from Group 1.

Table B-11. Item Difficulties (p), Point-Biserial Correlations (rpb), and Biserial Correlations (rbis), Within Session 1 for Version A of the EI Subtest by Form and Test Administration Medium

	Computer									Paper-and-Pencil								
	Form 11			Form 12			Form 13			Form 11			Form 12			Form 13		
Item	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis
1	.746	.375	.509	.870	.169	.269	.877	.247	.398	.772	.394	.546	.948	-.121	-.252	.897	.456	.772
2	.932	.175	.336	.722	.295	.394	.947	.489	1.016	.947	.074	.154	.862	.020	.032	.914	.371	.662
3	.780	.179	.250	.811	.182	.263	.857	.581	.901	.895	.332	.559	.879	.246	.399	.862	.395	.617
4	.741	-.009	-.012	.759	.318	.436	.772	.508	.705	.684	.175	.229	.724	.158	.211	.860	.160	.250
5	.862	-.061	-.095	.722	.284	.380	.732	.271	.364	.877	.116	.188	.759	.399	.547	.879	.259	.421
6	.797	.334	.475	.673	.241	.313	.750	.423	.576	.860	.244	.380	.914	.068	.122	.793	.124	.176
7	.763	.320	.440	.792	.235	.333	.709	.569	.754	.789	.176	.249	.737	.310	.418	.707	.362	.479
8	.741	.425	.575	.736	.395	.533	.737	.494	.666	.768	.590	.817	.793	.023	.033	.724	.439	.587
9	.793	.281	.398	.792	.361	.512	.772	.407	.564	.754	.407	.556	.603	.331	.420	.845	.182	.276
10	.763	.295	.407	.660	.499	.645	.754	.217	.296	.772	.275	.381	.719	.106	.142	.759	.309	.425
11	.621	.370	.472	.685	.414	.542	.772	.025	.035	.684	.651	.852	.741	.072	.098	.707	.140	.185
12	.508	.207	.259	.692	.451	.592	.491	.441	.552	.737	.172	.233	.741	.058	.079	.690	.314	.411
13	.655	.264	.341	.472	.340	.426	.754	.407	.557	.536	.233	.292	.603	.184	.233	.672	.116	.150
14	.441	.303	.381	.585	.102	.130	.632	.143	.183	.536	.533	.669	.672	.331	.430	.586	.196	.248
15	.525	.510	.639	.491	.575	.720	.544	.441	.554	.589	.327	.414	.638	.226	.290	.534	.420	.527
16	.492	.436	.547	.623	.377	.481	.684	.243	.317	.464	.427	.536	.534	.238	.299	.741	.322	.436
17	.552	-.043	-.054	.673	.330	.429	.464	.176	.220	.554	-.049	-.062	.483	.015	.019	.448	.191	.240
18	.305	.357	.469	.538	.462	.580	.643	.455	.584	.393	.435	.553	.517	.273	.342	.719	.190	.254
19	.492	.480	.602	.547	.075	.094	.339	-.027	-.035	.607	.445	.566	.509	.223	.280	.386	.319	.406
20	.288	.388	.514	.481	.320	.401	.263	.010	.013	.339	.421	.544	.421	.302	.381	.316	.227	.296

Note. Computer data are from Group 1 and paper-and-pencil data are from Group 2.

Table B-12. Item Difficulties (p), Point-Biserial Correlations (rpb), and Biserial Correlations (rbis), Within Session 1 for Version B of the EI Subtest by Form and Test Administration Medium

	Computer									Paper-and-Pencil								
	Form 11			Form 12			Form 13			Form 11			Form 12			Form 13		
Item	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis
1	.764	.100	.137	.875	.311	.499	.875	.446	.717	.821	.319	.467	.945	.143	.294	.879	.150	.243
2	.909	.061	.107	.786	.235	.331	.927	.440	.826	.911	.194	.342	.873	.158	.253	.983	.226	.689
3	.836	.357	.536	.839	.409	.616	.911	.448	.791	.964	.033	.078	.945	.143	.294	.931	.267	.509
4	.818	.034	.050	.589	.142	.179	.857	.252	.390	.714	.425	.565	.691	.387	.507	.810	.289	.418
5	.745	.370	.503	.709	.477	.631	.618	.347	.442	.821	.467	.685	.727	.399	.534	.810	.163	.236
6	.764	.438	.603	.764	.355	.489	.750	.382	.520	.804	.206	.295	.782	.366	.514	.810	.463	.669
7	.873	-.073	-.116	.764	.445	.613	.782	.471	.660	.911	.263	.464	.727	.184	.247	.879	.134	.217
8	.800	.309	.441	.696	.300	.395	.768	.105	.145	.857	.351	.544	.782	.258	.361	.672	.281	.365
9	.727	.469	.628	.714	.271	.360	.821	.224	.328	.750	.391	.533	.691	.276	.362	.690	.355	.466
10	.709	.432	.573	.732	.330	.444	.661	.366	.473	.786	.579	.815	.727	.258	.346	.828	.212	.314
11	.673	.316	.411	.661	.144	.186	.839	.346	.521	.661	.527	.682	.782	.312	.437	.828	.212	.314
12	.636	.155	.198	.625	.399	.510	.571	.280	.353	.607	.273	.347	.778	.153	.214	.638	.354	.453
13	.600	.064	.081	.741	.582	.788	.500	.066	.083	.500	.403	.505	.741	.214	.290	.741	.289	.391
14	.564	.531	.668	.750	.258	.351	.582	.089	.113	.564	.431	.543	.582	.043	.054	.655	.269	.347
15	.636	.410	.525	.661	.359	.465	.661	.204	.264	.625	.390	.498	.655	.454	.586	.707	.358	.474
16	.455	.297	.373	.600	.339	.429	.714	.172	.229	.518	.140	.176	.630	.276	.353	.793	.212	.301
17	.600	.468	.593	.607	.297	.378	.309	.169	.222	.679	.590	.770	.537	.412	.518	.431	.522	.657
18	.455	.135	.170	.589	.329	.417	.564	.244	.308	.418	.277	.350	.500	.377	.472	.690	.368	.482
19	.491	-.112	-.140	.518	.284	.356	.464	.297	.372	.455	.067	.085	.481	.247	.309	.569	.352	.443
20	.309	.394	.516	.455	.166	.208	.357	.090	.116	.283	.331	.441	.528	.321	.403	.293	.262	.347

Note. Computer data are from Group 1 and paper-and-pencil data are from Group 2.

Table B-13. Item Difficulties (p), Point-Biserial Correlations (rpb), and Biserial Correlations (rbis), Within Session 2 for Version A of the EI Subtest by Form and Test Administration Medium

	Computer									Paper-and-Pencil								
	Form 11			Form 12			Form 13			Form 11			Form 12			Form 13		
Item	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis
1	.842	.105	.159	.982	.368	1.102	.836	.362	.542	.825	.442	.651	.860	.172	.267	.930	.166	.315
2	.860	.269	.419	.800	.281	.401	.945	.092	.190	.912	.156	.278	.842	.243	.368	.912	.083	.147
3	.930	.175	.332	.857	.329	.510	.945	-.012	-.025	.807	.198	.286	.877	.175	.282	.912	.232	.413
4	.719	.362	.482	.679	.481	.627	.873	.128	.204	.860	.197	.307	.702	.079	.105	.912	.157	.279
5	.821	.398	.584	.719	.453	.604	.873	.411	.657	.807	.184	.266	.737	.271	.365	.842	.421	.636
6	.789	.322	.455	.696	.383	.504	.800	.378	.541	.789	.512	.724	.807	.421	.606	.877	.214	.346
7	.895	.231	.390	.745	.094	.128	.836	.275	.412	.804	.073	.105	.825	.428	.631	.649	.519	.668
8	.825	.582	.857	.842	.555	.840	.636	.197	.252	.860	.150	.234	.719	.311	.415	.649	.334	.430
9	.719	.385	.513	.684	.479	.626	.709	.216	.286	.842	.231	.349	.737	.197	.266	.825	.265	.390
10	.789	.501	.708	.745	.258	.350	.815	.205	.298	.789	.273	.386	.702	.195	.258	.772	.278	.386
11	.667	.338	.438	.679	.475	.620	.800	.313	.447	.789	.498	.703	.754	.294	.402	.842	.238	.359
12	.649	.303	.390	.714	.345	.458	.618	.373	.475	.649	.214	.275	.684	.272	.355	.632	.353	.452
13	.509	.382	.479	.714	.345	.458	.691	.297	.390	.579	.292	.369	.667	.208	.270	.684	.484	.632
14	.561	.435	.548	.579	.074	.093	.618	.277	.353	.509	.429	.538	.684	.356	.465	.571	.185	.234
15	.667	.382	.496	.684	.322	.420	.691	.213	.280	.526	.294	.369	.772	.399	.553	.544	.321	.404
16	.526	.171	.215	.579	.368	.464	.745	.269	.365	.526	.340	.427	.596	.331	.420	.737	.560	.755
17	.684	.409	.534	.614	.322	.409	.444	.327	.412	.404	-.078	-.099	.561	.238	.300	.509	.278	.349
18	.357	.388	.498	.464	.494	.620	.800	.170	.243	.421	-.002	-.003	.544	.162	.204	.614	.398	.506
19	.368	.131	.168	.439	.251	.316	.491	.275	.345	.561	.547	.689	.632	.409	.523	.526	.036	.046
20	.351	.100	.128	.421	.330	.417	.309	.020	.026	.351	.444	.572	.474	.300	.376	.286	.229	.304

Note. Computer data are from Group 2 and paper-and-pencil data are from Group 1.

Table B-14. Item Difficulties (p), Point-Biserial Correlations (rpb), and Biserial Correlations (rbis), Within Session 2 for Version B of the EI Subtest by Form and Test Administration Medium

	Computer									Paper-and-Pencil								
	Form 11			Form 12			Form 13			Form 11			Form 12			Form 13		
Item	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis	p	rpb	rbis
1	.786	.389	.548	.966	.086	.204	.893	.533	.893	.759	.302	.415	.842	.205	.310	.873	.263	.420
2	.946	.092	.190	.877	.055	.088	.946	.377	.779	.931	.348	.663	.825	.252	.371	.982	.215	.643
3	.929	.214	.405	.845	.142	.215	.821	.437	.641	.810	.454	.656	.877	.297	.479	.891	.200	.333
4	.750	.027	.037	.724	.221	.295	.857	.103	.159	.776	-.009	-.012	.789	.304	.429	.800	.490	.700
5	.839	.001	.002	.759	.221	.303	.821	.375	.550	.759	.265	.363	.807	.259	.373	.800	.286	.409
6	.857	.272	.422	.879	.173	.280	.714	.413	.550	.810	.247	.358	.719	.397	.529	.836	.447	.670
7	.821	.311	.457	.776	.267	.373	.750	.359	.489	.879	.149	.242	.789	.279	.394	.800	.271	.387
8	.750	.565	.769	.724	.364	.486	.732	.400	.537	.810	.274	.397	.789	.255	.360	.745	.502	.682
9	.786	.352	.495	.655	.366	.473	.857	-.022	-.034	.702	.415	.548	.772	.409	.567	.782	.167	.234
10	.768	.434	.600	.638	.044	.057	.750	.317	.433	.845	.367	.558	.667	.452	.585	.764	.099	.137
11	.786	.594	.836	.737	.165	.222	.750	.196	.268	.759	.340	.467	.825	.344	.506	.818	.259	.378
12	.786	.229	.322	.684	-.077	-.101	.714	.360	.478	.500	.281	.352	.667	.139	.180	.600	.145	.184
13	.500	.337	.422	.672	.336	.437	.661	.158	.205	.500	.403	.505	.649	.531	.684	.564	.352	.443
14	.625	.549	.702	.655	.298	.384	.582	.118	.149	.552	.287	.361	.561	.356	.449	.655	.340	.439
15	.589	.364	.461	.672	.254	.330	.607	.170	.215	.672	.312	.406	.482	.244	.307	.727	.397	.533
16	.464	.445	.559	.586	.163	.206	.732	.400	.537	.448	.382	.480	.632	.325	.416	.611	.202	.257
17	.643	.219	.281	.534	.089	.111	.411	.280	.354	.614	.415	.528	.579	.216	.273	.327	.081	.105
18	.411	.361	.457	.534	.174	.219	.714	.178	.237	.333	.243	.315	.589	.528	.668	.727	.355	.475
19	.643	.445	.571	.561	.314	.396	.393	.084	.107	.500	-.169	-.212	.632	.210	.268	.382	.189	.241
20	.393	.497	.631	.386	.048	.062	.357	.200	.257	.357	.289	.372	.446	.216	.272	.345	.027	.034

Note. Computer data are from Group 2 and paper-and-pencil data are from Group 1.