AFHRL-TR-71-42

AIR FORCE

HUMAN RESOURCES

LABORATORY

AD734740

A PERFORMANCE EVALUATION: MICROFICHE VERSUS HARDCOPY

By

Robert R. Grausnick
James P. Kottenstette

Denver Research Institute
University of Denver

TECHNICAL TRAINING DIVISION
Lowry Air Force Base, Colorado

May 1971

DDC
RECEIVED
NOV 29 1971
B

Approved for public release; distribution unlimited.

Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
Springfield, Va. 22151

AIR FORCE SYSTEMS COMMAND

BROOKS AIR FORCE BASE, TEXAS

APPS SA

62

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*
University of Denver
Denver Research Institute
University Park, Denver Co 80210

2a. REPORT SECURITY CLASSIFICATION

2b. GROUP

3. REPORT TITLE

A PERFORMANCE EVALUATION: MICROFICHE VERSUS HARDCOPY

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*
Final Report (July 1970 to July 1971)

5. AUTHOR(S) *(First name, middle initial, last name)*
Robert R. Grausnick
James P. Kottenstette

6. REPORT DATE
May 1971

7a. TOTAL NO. OF PAGES
51

7b. NO. OF REFS
9

8a. CONTRACT OR GRANT NO.
F-41609-70-C-0040

b. PROJECT NO. 1121

c. Task No. 112101

d. Work Unit No. 112101009

9a. ORIGINATOR'S REPORT NUMBER(S)
AFHRL-TR-71-42

9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*

10. DISTRIBUTION STATEMENT
Approved for public release; distribution unlimited.

11. SUPPLEMENTARY NOTES

12. SPONSORING MILITARY ACTIVITY
Technical Training Division
Air Force Human Resources Laboratory
Lowry AFB CO 80230

13. ABSTRACT

The primary purpose of the ~~present~~ study was to replicate an experiment previously conducted by Baldwin and Bailey (1971) in order to establish the reproducibility of their results in a different user environment. Twelve psychometric instruments employing technical training materials which required various types of visual skills were used in the testing. Three forms of each test exercise were reproduced: hardcopy, positive-image microfiche, and negative-image microfiche. Ninety subjects were randomly assigned to one of three experimental groups (hardcopy, positive-image, and negative-image). An analysis of variance (Winer, 1962; Myers, 1966) revealed F values to be significant beyond the .01 level for Test 4, Figure Identification; Test 6, Symbol Translation; Test 7, Graphs; Test 8, Tables; and Test 12, Number Verification; and beyond the .05 level for Test 5, Length Estimation. No significant differences in performance were found between positive and negative-image microfiche presentations. These results generally substantiate those reported by Baldwin and Bailey. Further analyses demonstrated that the reader presentation primarily affected the speed at which the subjects worked while accuracy was not differentially affected by the presentation mode. An analysis of the performance of subjects grouped according to Armed Forces Qualification Test (AFQT) scores indicated that various intelligence groups were differentially affected by the mode of presentation. While statistical differences were found in a number of test exercises, no fundamental difficulties were encountered which would bar future utilization of microform materials in technical training programs.

DD FORM 1 NOV 65 1473

Unclassified
Security Classification

NOTICE

When US Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

ACCESSION for
CFSTI WHITE SECTION ☑
DDC BUFF SECTION ☐
UNANNOUNCED ☐
JUSTIFICATION .....
BY .....
DISTRIBUTION/AVAILABILITY CODES
DIST. AVAIL. and/or SPECIAL
A

14 KEY WORDS	LINK A ROLE	LINK A WT	LINK B ROLE	LINK B WT	LINK C ROLE	LINK C WT
Psychometric instruments						
readability						
visibility						
microfiche						

ABSTRACT

The primary purpose of the present study was to replicate an experiment previously conducted by Baldwin and Bailey (1971) in order to establish the reproducibility of their results in a different user environment. Twelve psychometric instruments employing technical training materials which required various types of visual skills were used in the testing. Three forms of each test exercise were reproduced: hardcopy, positive-image microfiche, and negative-image microfiche. Ninety subjects were randomly assigned to one of three experimental groups (hardcopy, positive-image, and negative-image). An analysis of variance (Winer, 1962; Myers, 1966) revealed F values to be significant beyond the .01 level for Test 4, Figure Identification; Test 6, Symbol Translation; Test 7, Graphs; Test 8, Tables; and Test 12, Number Verification; and beyond the .05 level for Test 5, Length Estimation. No significant differences in performance were found between positive and negative-image microfiche presentations. These results generally substantiate those reported by Baldwin and Bailey. Further analyses demonstrated that the reader presentation primarily affected the speed at which the subjects worked while accuracy was not differentially affected by the presentation mode. An analysis of the performance of subjects grouped according to Armed Forces Qualification Test (AFQT) scores indicated that various intelligence groups were differentially affected by the mode of presentation. While statistical differences were found in a number of test exercises, no fundamental difficulties were encountered which would bar future utilization of microform materials in technical training programs.

AFHRL-TR-71-42

May 1971

A PERFORMANCE EVALUATION: MICROFICHE VERSUS HARDCOPY

By

Robert R. Grausnick
James P. Kottenstette

Denver Research Institute
University of Denver

Approved for public release; distribution unlimited.

TECHNICAL TRAINING DIVISION
AIR FORCE HUMAN RESOURCES LABORATORY
AIR FORCE SYSTEMS COMMAND
Lowry Air Force Base, Colorado

FOREWORD

The research to be reported was conducted by the Denver Research Institute, University of Denver, under Contract No. F41609-70-C-0040. Mr. James P. Kottenstette was the Principal Investigator and Mr. Robert R. Grausnick was Research Associate. The contract as a whole extends from July 1, 1970 to July 1, 1971. The research reported here, however, was conducted from August 1, 1970 to October 10, 1970. This study represents a portion of the research program of Project 1121, Technical Training Development; Task 112101, Advanced Technology for Air Force Technical Training. Dr. Marty R. Rockway was the Project Scientist and Mr. Joseph Y. Yasutake was the Task Scientist.

The authors wish to acknowledge the cooperation and support of the Contract Monitor, Dr. Edgar A. Smith, Air Force Human Resources Laboratory, Lowry Air Force Base, Colorado. Acknowledgement also is made to Captain John J. Zucarrelli and Sergeant Minor P. Johnson of Lowry Air Force Base for their assistance in providing the subjects used in the experimentation. Appreciation is extended to Mr. Jerry Siener for his assistance in the collection of data and to Miss Anne Dailey for her help with the statistical analyses.

Other reports prepared under this contract include:

AFHRL-TR-71-43, "Microform Use in a Technical Training Environment--An Experiment"

AFHRL-TR-71-44, "A Guide to Instructional Uses of Microform"

This report has been reviewed and is approved.

GEORGE K. PATTERSON, Colonel, USAF
Commander

SUMMARY

Denver Research Institute, University of Denver. A Performance Evaluation: Microfiche versus Hardcopy. AFHRL-TR-71-42, Lowry AFB, Colo.: Technical Training Division, Air Force Human Resources Laboratory, May 1971.

Objectives

The primary objectives of this study were (a) to replicate an experiment conducted by Baldwin and Bailey in 1971 to establish the reproducibility of their results in a different user environment, and (b) to establish the reliability of instruments which they developed as tools for use in future Air Force microfiche evaluations.

Approach

Three forms (hardcopy, positive-image microfiche, and negative-image microfiche) of twelve psychometric instruments employing technical training materials were developed for this study. Ninety Air Force trainees were randomly assigned to one of three experimental groups (hardcopy, positive-image and negative-image) and performance comparisons were evaluated in terms of the study objectives.

Findings

The major findings in terms of the replicative aspects of the study were:

1. Baldwin and Bailey's finding of no significant difference in performance, using narrative-type materials, among any of the presentation modes was confirmed in terms of the overall experimental results.

2. The original finding of no significant difference in performance across all instruments for the positive versus the negative microfiche presentations was also confirmed.

3. Significant differences between fiche and hardcopy performance for three of the nine non-narrative, visual discrimination instruments were encountered in both studies, but three additional instruments

also showed significant performance differences between hardcopy and one or both of the fiche presentations in the present study.

In general, then, the Baldwin and Bailey results were replicated; their main results could be extended to a different environment, a different time, and to different equipment. Further analyses resulted in two additional findings:

1. The microform presentation primarily affected the speed at which the subjects worked while accuracy was not differentially affected by the presentation.

2. An analysis of the performance of subjects grouped according to AFQT scores indicated that various intelligence troups were differentially affected by the mode of presentation.

This summary was prepared by Dr. Edgar A. Smith, Technical Training Division, Air Force Human Resources Laboratory .

TABLE OF CONTENTS

		Page
I.	INTRODUCTION	1
	Purpose of Study	3
	Hypotheses	4
II.	METHOD	6
	Psychometric Instruments	6
	Equipment and Testing Environment	10
	Subjects	10
	Testing Procedures	11
III.	RESULTS	12
IV.	DISCUSSION AND CONCLUSIONS	25
REFERENCES		33
APPENDIX A.	Test Instructions and Sample Items	34
APPENDIX B.	Orders of Presentation	45
APPENDIX C.	Computing the Scheffe Method of Post-hoc Comparisons	50
APPENDIX D.	Test Means and Standard Deviations in the Baldwin and Bailey Study	51

LIST OF TABLES

Table No.		
I.	Reliability Coefficients	9
II.	Analysis of Variance for 12 Tests	13
III.	Test Means and Standard Deviations	15
IV.	Scheffe Values for the Twelve Tests	16
V.	Percentage of Attempted Responses Which Were Correct	18

Table No.		Page
VI.	Test Means and Standard Deviations for AFQT Scores 80-100	19
VII.	Test Means and Standard Deviations for AFQT Scores 50-79	20
VIII.	Test Means and Standard Deviations for AFQT Scores 0-49	21
IX.	Scheffe Values for the Twelve Tests (80-100 AFQT Scores)	23
X.	Scheffe Values for the Twelve Tests (0-49 AFQT Scores)	24

A PERFORMANCE EVALUATION: MICROFICHE VERSUS HARDCOPY

I. INTRODUCTION

Microforms have long been used in industry and government as a storage and retrieval tool having enormous administrative value and varying user value.* In educational environments, however, microforms (although now widely distributed) reflect primarily an administrative solution to the problems of acquiring and storing specialized materials. This results in a limited number of applications and, therefore, a limited user group. The ability of individuals to effectively use microform training materials must be examined before educational applications can be expanded from their substitution for hardcopy to their routine utilization as a new communication medium.

Experimental studies are needed which consider the question, "To what extent are the cognitive skill levels of students preserved when using a microform presentation?" In one such study, Kottenstette (1969) determined in a reading experiment that there are no fundamental physical or psychological barriers to the utilization of microforms in the communication of narrative information that the student customarily encounters in hardcopy. Students are able to preserve skill levels (reading rate and comprehension) when utilizing reader presentations of both descriptive and abstract narrative materials which reflect various levels of difficulty.

* The term microform, as employed in this report, is intended to include microfiche, ultrafiche, ultra-reduced microcopy and photochromatic images. The COSATI standard microfiche form was used in this study.

Baldwin and Bailey (1971) also found that, for narrative reading at least, students could perform adequately using microfilmed materials. These investigators conducted an extensive and thorough review of the technical training materials used at the Chanute Technical Training Center, Chanute Air Force Base, Illinois. After consulting with a number of experienced training specialists, they developed twelve tests which represent the various visual skills involved in the utilization of these materials.* Three forms of each test master were reproduced (hardcopy, positive-image microfiche, and negative-image microfiche) and presented to three separate trainee groups. The experimental results indicated statistically significant differences favoring the hardcopy presentation for three of the twelve tests. All of the significant differences were encountered in tests which involved character recognition and symbol interpretation as opposed to reading continuous prose. The authors concluded that, in terms of "readability," materials presented via microfiche are feasible for future technical training purposes.

"Readability," as conventionally used in relation to hardcopy, has been defined in terms of the difficulty or complexity of the content of the material. Applying this definition, more readable (easier, less complex) material is likely to result in greater understanding, learning, and retention than less readable (more difficult, more complex) material (Klare, 1963). When hardcopy materials are transferred to microfiche, a somewhat different definition is employed which can best be understood in terms of the distinction made by Kottenstette (1969) between readability and visibility.

---

* It should be emphasized that the tests developed were not designed to represent technical training content areas but rather visual skills.

Some of the image quality of the hardcopy is lost in the transfer to microfiche, both in the photographic reduction process and in the reader magnification process. A substantial loss of image quality can be tolerated with little effect on the user's reading skills when readability depends on the recognition of complete words or groups of words in context, as in narrative materials. However, with materials which, by their nature, content, or purpose are dependent upon individual character recognition, or visibility, the reduced image quality may affect cognitive skill levels.

The distinction made here between readability and visibility is important because the tests developed by Baldwin and Bailey focused on the visual skills utilized in training, not on the content of the materials encountered. This suggests that the successful completion of many of the test exercises was more dependent upon visibility than upon readability, as these terms are last defined above. However, in some cases, the exercises involved more than a simple visual discrimination task; in order to complete the exercise, the subjects were required to assimilate, process, and use that which was discriminated. In this way, the term "readability" was expanded in the Baldwin and Bailey study to include the interpretation and coding skills as well as the visual discrimination skills involved in dealing with schematics, wiring diagrams, flow charts, tables, graphs, and three-dimensional drawings.

Purpose of Study

The present study is part of a larger research effort designed to develop and evaluate microforms for use as primary source materials in Air Force technical training programs. Data from three

investigations will be analyzed and reported separately in response to the following questions:

(1) Can Air Force trainees use microform presentations of educational and training materials to an extent consistent with their use of hardcopy materials?

(2) What are the optimal microform formats to be used in presenting technical training materials?

(3) What are the costs, advantages, and disadvantages of microform presentations in actual classroom use?

In addition, a user-oriented guide to the utilization of microform technology in technical training will be prepared, based on a review of the literature and the insights which result from the research effort.

The research reported here addresses the first of the questions enumerated above. It was designed primarily to replicate the experiment conducted by Baldwin and Bailey in order to establish the reproducibility of their results in a different user environment, and to establish the reliability of instruments which they developed as tools for use in future Air Force microfiche evaluations. This study had the secondary purposes of providing insights into the strengths and weaknesses of microform presentations in an operational sense, as well as generating criteria for evaluating the psychological or training implications of using microforms as primary source materials in technical training.

Hypotheses

Two hypotheses were developed to help meet the primary research objective. The first stated that (1) students using the hardcopy presentation would perform significantly better than students using

either the positive-image or the negative-image microfiche presentations. This hypothesis was based on Baldwin and Bailey's finding that when the method of presentation affects the "readability" of material, the advantage favors the hardcopy presentation. The second hypothesis stated that (2) there would be no significant performance difference between subjects using positive-image microfiche and subjects using negative-image microfiche. The second hypothesis was stated in a non-directional form because the question of film polarity had not been adequately resolved before the present study was conducted.

Baldwin and Bailey encountered enormous variance in their data suggesting that a stratification of the subjects based on an intelligence index might prove valuable. It appeared possible that highly intelligent subjects would respond to the machine presentation with increased or decreased sensitivity as compared with less intelligent subjects. An additional hypothesis, therefore, stated that (3) subjects divided into high, medium, and low intellectual groups would be differentially affected by the presentation mode.

II. METHOD

Psychometric Instruments

The test exercises used in this study were developed by Baldwin and Bailey and required many of the types of visual skills used in Air Force technical training programs. Performance requirements included: the reading of continuous prose and short paragraphs; reading schematics, wiring diagrams, and flow charts; identifying and discriminating precise figures and symbols; and reading charts, tables, graphs, and three-dimensional drawings.

Several of these test instruments were reproduced or adapted from the Kit of Reference Tests for Cognitive Factors, copyrighted by the Educational Testing Service, Princeton, New Jersey, and developed under Office of Naval Research Contract Nonr-2214(00), Project Designation NR-151-174 (1963). Reproduction of these test instruments is permitted by and for United States Government use.

The following paragraphs contain a short description of each test exercise. The time limits presented were taken from Baldwin and Bailey who established them at two standard deviations above the mean time for completion of one-half of the test (based on pretest data; N = 42). All tests were scored by counting the number of correct responses for each test exercise. The test instructions and sample items of each test are included as Appendix A.

Tests 1 and 2, Narrative-Continuous Prose. Material for this section was selected from the survival training text of the ATC Student Study Guide OZR 1515A. The content was rewritten as continuous, uninterrupted prose and modified by deleting three key words on each page and replacing them with blanks. The task was to read the

material and supply the missing words (Part I 50 minutes; Part II 25 minutes).

Test 3, Narrative-Short Paragraph. The same type of material as used in the continuous prose exercises was used in this test; one-hundred paragraphs, two or three sentences in length, were isolated from the text and modified by changing one word in each paragraph to be inconsistent with the meaning of that paragraph. The subjects were required to read the paragraph and identify the inconsistent word (25 minutes).

Test 4, Figure Identification. For each item, the subjects indicated which of five geometrical figures or pictures in a row was identical to a figure presented at the left of the row (5 minutes).

Test 5, Length Estimation. Each item consisted of two points connected by three curved or angular lines. The task of the subjects was to select the shortest of the lines (3 minutes).

Test 6, Symbol Translation. This was a test of the subjects' ability to translate symbols into alphabetic characters. A legend presented alphabetic letters and their corresponding symbols. Each item consisted of a short series of symbols; the subjects were required to identify the alphabetic characters which corresponded to each symbol (10 minutes).

Test 7, Graphs. Four curves on a graph plotted the relationship of temperature to time (in minutes) under certain experimental conditions. For each test item, the minutes were given. The task was

to indicate the temperature which corresponded to the given number of minutes for each curve (7 minutes).

Test 8, Tables. A table was presented which contained drill-bit sizes and their corresponding decimal equivalents. The task was to refer to the table and indicate the decimal equivalent for each of the given drill-bit sizes (4 minutes).

Test 9, Mechanical Drawing. Subjects were required to locate lines and surfaces on three separate views of an object (front, side, and top) corresponding to lines and surfaces on a three-dimensional view of the same object (5 minutes).

Test 10, Schematics. An item consisted of a network of lines as in an electrical-current diagram which has many intersecting and intermeshing wires and five sets of terminals, each marked S (start) and F (finish). The task was to follow the lines and to determine through which pair of terminals there was a complete circuit from S, through a circle at the top of the diagram, to F (15 minutes).

Test 11, Flow Diagrams. The subjects were presented with diagramatic sections representing city maps with the streets blocked at various points by barriers. They were required to select the shortest path between two points so that no roadblocks need be crossed (7 minutes).

Test 12, Number Verification. The subjects inspected pairs of multi-digit numbers and indicated whether the two sets of numbers differed or were identical (6 minutes).

Baldwin and Bailey determined the reliability of each test using the following procedure: One-half of each test was administered on successive days to a sample of 42 subjects, using only the hardcopy presentation. After test reliabilities were computed, the Spearman-Brown prophecy formula was used to estimate reliabilities for the total test length. Results of the analysis are presented in Table I and indicate that, in general, the reliabilities are within a range acceptable for making comparisons among treatment groups. A possible exception is Test 11, Flow Diagrams, which has a reliability of .59. It would be necessary to increase this exercise to 2.78 times its present length to achieve a reliability of .80. The remaining eleven tests have reliabilities ranging from .76 to .93.

Table I. Reliability Coefficients

N = 42

Test No.		r_{xx}
1.	Narrative (50 minutes)	.92
2.	Narrative (25 minutes)	.85
3.	Narrative - Short Paragraph	.79
4.	Figure Identification	.90
5.	Length Estimation	.90
6.	Symbol Translation	.79
7.	Graphs	.93
8.	Tables	.76
9.	Mechanical Drawing	.90
10.	Schematics	.88
11.	Flow Diagrams	.59
12.	Number Verification	.91

The test masters prepared by Baldwin and Bailey were used to reproduce the testing instruments used in this experiment. It should be emphasized that the same masters were used to produce the hardcopy and to film the positive-image and negative-image microfiche forms.

Equipment and Testing Environment

The reader used in this study, the Eastman Kodak Recordak Easamatic Reader, Model PFCD, is a tabletop, film reader designed specifically for viewing images on 4- by 6-inch microfiche having DOD, NMA or COSATI formats. The readers could accommodate both positive- and negative-image microfiche.

The experiment was conducted at the Human Resources Laboratory, Lowry Air Force Base, Colorado. A large room was provided and equipped with six 30- by 54-inch tables. Four tables were used to accommodate microfiche readers and the remaining two were used for hardcopy work.

The ambient illumination level normally varied from approximately 40 foot-candela to 50 foot-candela over the duration of the daily experimental period. The variation, due to sun position, was well within the comfort range for reading offset copy. The overall environment was typical of an office setting.

Subjects

Experimental subjects were obtained from Air Force technical training students or Personnel Awaiting Training Status (PATS) at Lowry Air Force Base, Colorado. All subjects were high school graduates; none held college degrees, although some had limited college experience. To test the hypothesis of no difference among the subject distributions in the three treatment groups used, a chi-square test using the Armed Forces Qualification Test (AFQT) scores was computed.

The AFQT includes vocabulary, arithmetic reasoning, and spatial relationship problems, the last-named involving the recognition, perception, manipulation, and analysis of relations in two and three dimensions (Uhlaner, 1952). Items were selected on the basis of difficulty level as well as on the basis of their correlations with subtests and the total test scores.

The resultant chi square value of 9.835 for 18 degrees of freedom was not significant ($p > .50$), and indicated that there were no significant differences among the three groups in terms of AFQT scores.

Testing Procedures

Ninety subjects were tested, 30 in each of the three treatment groups.* All subjects were airmen awaiting training at the Lowry Technical Training Center, Lowry Air Force Base, Colorado. The subjects were assigned in a quasi-random manner to testing positions resulting in two subjects for each of the three treatment groups each session. All test instruments were administered by the same investigator using standardized test procedures and instructions. The total test battery required approximately four hours to complete. Forty minutes of this time, however, was devoted to breaks; two ten-minute breaks and one twenty-minute break each day. The order of presentation was counterbalanced to distribute possible practice and fatigue effects equally over all conditions. That is, the instruments were administered in fifteen different sequences within each group, one for each test day. The orders of presentation are presented in Appendix B.

* 94 students were actually run: four subjects were eliminated because they were unable to complete the entire test sequence.

III. RESULTS

The primary purpose of this study was to replicate a previous experiment (Baldwin and Bailey) to determine the effect of three different methods of presentation on the ability of subjects to process several types of information through the visual modality. As in the original experiment, separate analyses of variance were computed for each of the twelve test exercises.

The results of the single classification analyses of variance are presented in Table II. F values for the test exercises were found to be significant beyond the .01 level for Test 4, Figure Identification; Test 6, Symbol Translation; Test 7, Graphs; Test 8, Tables; and Test 12, Number Verification; and beyond the .05 level for Test 5, Length Estimation. (These exercises include Tests 4, 6, and 7, Figure Identification, Symbol Translation and Graphs, in which significance was also found in the original study by Baldwin and Bailey using the same procedure.)

Since the analyses of variance resulted in significant F values for Tests 4, 5, 6, 7, 8, and 12, further analyses were conducted to determine the significance of differences between pairs of means. The Scheffe method of making post-hoc comparisons (Scheffe, 1959) was selected for use in this study because it is applicable in situations where a preliminary analysis of variance has shown overall significance. In addition, the Scheffe test has no requirement that post-hoc comparisons be independent; it can be used to make any and all comparisons of interest to the investigator. It was used, therefore, to compare all pairs of means in the analysis. The Scheffe method is more conservative than other multiple comparison methods with regard to Type-I error, and leads to fewer significant differences. Since the Scheffe method is so conservative, this study followed the

Table II. Analysis of Variance for 12 Tests

Test	Source of Variation	Sum of Squares	df	Mean Square	F Ratio
Narrative (50)	Between groups	668.42	2	334.21	.588
	Within groups	49407.90	87	567.91	
	Total	50076.32	89	562.66	
Narrative (25)	Between groups	310.40	2	155.20	.828
	Within groups	16314.50	87	187.52	
	Total	16624.90	89	186.80	
Short Paragraph	Between groups	399.09	2	199.54	1.205
	Within groups	14404.03	87	165.56	
	Total	14803.12	89	166.33	
Figure Identification	Between groups	2895.62	2	1447.81	14.083**
	Within groups	8944.33	87	102.81	
	Total	11839.96	89	133.03	
Length Estimation	Between groups	422.49	2	211.24	4.452*
	Within groups	4127.83	87	47.45	
	Total	4550.32	89	51.13	
Symbol Translation	Between groups	20376.80	2	10188.40	7.996**
	Within groups	110856.80	87	1274.22	
	Total	131233.60	89	1474.53	
Graphs	Between groups	1197.07	2	598.53	6.040**
	Within groups	8620.93	87	99.09	
	Total	9818.00	89	110.31	
Tables	Between groups	442.96	2	221.48	6.838**
	Within groups	2818.03	87	32.39	
	Total	3260.99	89	36.64	
Mechanical Drawing	Between groups	1059.36	2	529.68	1.318
	Within groups	34957.13	87	401.81	
	Total	36016.49	89	404.68	
Schematics	Between groups	1.87	2	0.93	0.026
	Within groups	3073.73	87	35.33	
	Total	3075.60	89	34.56	
Flow Diagrams	Between groups	138.29	2	69.14	1.630
	Within groups	3690.17	87	42.42	
	Total	3828.46	89	43.02	
Number Verification	Between groups	524.47	2	262.23	5.111**
	Within groups	4463.63	37	51.31	
	Total	4988.10	89	56.05	

* Significance beyond the .05 level
** Significance beyond the .01 level

recommendation of Ferguson (1966) that a less rigorous level of significance (.10) be used. A brief description of the computations involved in applying the Scheffe test are presented in Appendix C.

The means and standard deviations for each group, on each test, are presented in Table III. The results of the Scheffe tests for significance of differences between these means are presented in Table IV, and indicate that subjects using the hardcopy presentation performed significantly better than subjects using either the positive-image or the negative-image microfiche presentations for Test 4, Figure Identification; Test 5, Length Estimation; Test 6, Symbol Translation; and Test 12, Number Verification. In addition, test means for the hardcopy group were significantly higher than test means for the negative-image group only on Test 7, Graphs; and Test 8, Tables. These differences were significant at the suggested .10 level. None of the differences between means for the positive-image versus the negative-image microfiche groups were significant.

In the original experiment, Baldwin and Bailey, using t-tests to determine significant inter-group differences, found the mean score for hardcopy was significantly higher than the mean score for either positive or negative microfiche for Figure Identification, Symbol Translation, and Graphs.

The methodology established for the scoring of the test exercises (i.e., counting the number of correct responses made during established time limits) did not allow the consideration of speed and accuracy as separate factors. A post-hoc analysis, therefore, was conducted to provide some insights into this question. In this analysis, the ratio of correct responses to attempted responses was computed for each treatment group, on each test, as a rough measure of accuracy.

Table III. Test Means and Standard Deviations

Test	Hardcopy (N=30) $\bar{X}$	Hardcopy (N=30) S.D.	Positive (N=30) $\bar{X}$	Positive (N=30) S.D.	Negative (N=30) $\bar{X}$	Negative (N=30) S.D.
1. Narrative (50 min.)	69.23	24.30	64.67	20.85	62.73	24.94
2. Narrative (25 min.)	35.83	14.57	32.63	11.98	31.43	13.71
3. Short Paragraph	41.20	10.58	41.27	14.65	36.77	12.39
* 4. Figure Identification	59.80	11.10	49.83	10.12	46.43	8.52
* 5. Length Estimation	25.50	7.19	20.63	4.67	21.23	8.01
* 6. Symbol Translation	153.80	42.33	124.80	29.84	119.60	31.83
* 7. Graphs	38.80	10.01	34.33	10.59	29.87	8.66
* 8. Tables	18.70	5.92	16.07	5.60	13.27	5.25
9. Mechanical Drawing	33.67	21.36	41.17	18.85	40.70	18.80
10. Schematics	11.87	5.55	12.13	5.20	12.20	6.68
11. Flow Diagrams	20.07	7.35	19.50	5.86	17.20	5.89
*12. Number Verification	33.17	7.33	27.80	6.69	28.33	7.09

* Tests resulting in significant F-values (see Table II and text)

Table IV. Scheffe Values for the Twelve Tests

	Test	Hardcopy versus Positive	Hardcopy versus Negative	Positive versus Negative
1.	Narrative (50)	.4590	1.115	.100
2.	Narrative (25)	.8190	1.548	.120
3.	Short Paragraph	.0007	1.777	1.830
4.	Figure Identification	14.5000*	26.080*	1.680
5.	Length Estimation	7.4800*	5.750*	.110
6.	Symbol Translation	9.9000*	13.770*	.310
7.	Graphs	3.0200	12.070*	3.020
8.	Tables	3.2000	9.080*	3.630
9.	Mechanical Drawing	2.1000	1.840	.008
10.	Schematics	.0287	.003	.002
11.	Flow Diagrams	.1150	1.015	1.870
12.	Number Verification	8.4300*	6.850*	.080

* Significant beyond the .10 level (4.47)

The results, presented as percentages in Table V, indicate that students in all experimental groups were extremely accurate on all but three tests (Short Paragraph, Length Estimation and Schematics), and in these three tests, the groups were mutually consistent. This indicates that the test instruments were primarily speed tests; that is, any answer given had a high probability of being correct. Therefore, the significant differences in performance indicated in Table II were due to differences in the number of responses attempted, or speed effects. This interpretation is supported by the fact that there were no real differences among the three groups in terms of accuracy (as defined above); for any of the tests.

To test the hypothesis that students of varying intelligence are differentially affected by the mode of presentation, subjects were arbitrarily placed into the following groups based on their AFQT percentile scores: Group I, 80-100; Group II, 50-79; and Group III, 0-49.* Twelve one-way analyses of variance were computed for each of these groups. The resultant F ratios, as well as the means and standard deviations for the test exercises in each of the three AFQT groups, are presented in Tables VI, VII, and VIII. The analyses indicate that for Group I, F ratios were significant beyond the .05 level of significance for Test 6, Symbol Translation, and Test 7, Graphs; and beyond the .01 level for Test 4, Figure Identification.

In Group III, F ratios were found to be significant beyond the .05 level for Test 2, Narrative (25-minute), and Test 3, Short Paragraph; and beyond the .01 level for Test 5, Length Estimation;

* AFQT scores were unavailable for four subjects, resulting in an overall sample size of 86 for this analysis.

Table V. Percentage of Attempted Responses Which Were Correct

	Narrative (50)	Narrative (25)	Short Paragraph	Figure Ident.	Length Estima.	Symbol Trans.	Graphs	Tables	Mechanical Drawing	Schematic	Flow Diagrams	Number Verifica.
Hardcopy (N = 30)	94%	88%	83%	99%	76%	99%	93%	97%	90%	74%	92%	96%
Positive (N = 30)	93%	92%	84%	99%	70%	99%	95%	94%	97%	72%	92%	96%
Negative (N = 30)	94%	8[illegible]%	83%	98%	77%	98%	92%	94%	96%	78%	88%	94%

Table VI. Test Means and Standard Deviations for AFQT Scores 80-100

Test	Hardcopy (N=11) $\bar{X}$	Hardcopy (N=11) S.D.	Positive (N=12) $\bar{X}$	Positive (N=12) S.D.	Negative (N=14) $\bar{X}$	Negative (N=14) S.D.	F
1. Narrative (50 min.)	81.55	33.22	71.58	15.75	73.29	30.11	.398
2. Narrative (25 min.)	42.27	19.00	37.08	10.76	34.07	13.99	.876
3. Short Paragraph	44.64	12.40	47.92	11.83	44.71	11.15	.283
4. Figure Identification	62.18	9.96	50.92	11.43	45.57	8.67	7.931***
5. Length Estimation	26.18	6.38	21.00	4.85	22.00	9.95	1.391
6. Symbol Translation	160.36	45.43	131.50	33.77	115.93	29.40	4.309**
7. Graphs	43.09	8.58	38.50	9.91	33.14	7.91	3.665**
8. Tables	19.91	6.67	19.08	6.08	14.93	5.75	2.243
9. Mechanical Drawing	51.45	14.39	47.09	15.78	48.86	14.65	.227
10. Schematics	15.36	4.79	14.92	5.82	15.21	6.06	.017
11. Flow Diagrams	24.09	8.20	20.17	6.36	16.93	6.72	2.893*
12. Number Verification	32.55	5.65	29.17	7.91	30.14	7.08	.651

* Significant beyond the .10 level (2.47)
** Significant beyond the .05 level (3.32)
*** Significant beyond the .01 level (5.39)

Table VII. Test Means and Standard Deviations for AFQT Scores 50-79

Test	Hardcopy (N=10) $\bar{X}$	Hardcopy (N=10) S.D.	Positive (N=13) $\bar{X}$	Positive (N=13) S.D.	Negative (N=8) $\bar{X}$	Negative (N=8) S.D.	F*
1. Narrative (50 min.)	63.40	11.32	63.62	23.39	59.00	14.47	.169
2. Narrative (25 min.)	31.10	10.85	30.85	11.92	35.38	14.68	.343
3. Short Paragraph	40.30	9.37	38.08	13.26	33.50	3.32	.906
4. Figure Identification	57.80	9.75	51.08	6.87	48.88	9.87	2.458
5. Length Estimation	21.40	6.12	21.54	4.36	23.63	6.14	.416
6. Symbol Translation	145.10	35.48	125.31	26.00	131.13	32.30	1.065
7. Graphs	34.20	10.87	34.23	7.52	32.50	7.05	.106
8. Tables	16.90	4.32	14.54	4.34	14.50	3.91	.967
9. Mechanical Drawing	27.40	20.32	40.38	18.30	43.63	15.74	1.901
10. Schematics	9.60	3.41	10.62	3.79	11.38	4.64	.424
11. Flow Diagrams	19.90	5.05	20.77	5.09	19.00	4.33	.298
12. Number Verification	31.80	6.49	26.85	5.25	27.13	4.88	2.303

* None of the F-values were significant beyond the .10 level

Table VIII. Test Means and Standard Deviations for AFQT Scores 0-49

Test	Hardcopy (N=8) $\bar{X}$	Hardcopy (N=8) S.D.	Positive (N=3) $\bar{X}$	Positive (N=3) S.D.	Negative (N=7) $\bar{X}$	Negative (N=7) S.D.	F
1. Narrative (50 min.)	57.50	10.59	40.00	12.57	47.71	11.22	2.582
2. Narrative (25 min.)	31.75	6.38	20.33	10.27	21.71	6.43	3.940*
3. Short Paragraph	38.00	8.37	19.67	3.68	25.14	10.99	5.175*
4. Figure Identification	55.75	9.58	48.67	7.85	46.43	5.78	2.234
5. Length Estimation	28.25	6.74	15.67	1.25	17.43	3.16	9.918**
6. Symbol Translation	146.63	38.07	94.67	4.03	117.43	33.83	2.616
7. Graphs	36.50	5.72	17.33	10.96	20.71	5.03	11.984**
8. Tables	18.75	6.20	10.67	2.49	9.14	2.75	7.459**
9. Mechanical Drawing	20.50	12.51	19.33	15.20	18.29	10.22	.051
10. Schematics	9.63	6.36	8.33	2.36	5.43	2.92	1.253
11. Flow Diagrams	15.00	5.45	12.67	3.09	15.43	5.42	.266
12. Number Verification	34.75	9.59	24.67	6.18	26.43	8.68	1.930

* Significant beyond the .05 level (3.68)
** Significant beyond the .01 level (6.36)

Test 7, Graphs; and Test 8, Tables. There were no significant differences for Group II.

Since significant differences were found in both the 80-100 percentile groups and the 0-49 percentile groups, a Scheffe test of post-hoc comparisons was computed for all pairs of means in these groups. The results, presented in Tables IX and X, indicate that for Group III (0-49 AFQT percentile scores) the mean score for subjects using the hardcopy presentation were significantly higher than the mean scores for subjects using either the positive-image or the negative-image fiche presentations for Test 3, Short Paragraph; Test 5, Length Estimation; Test 7, Graphs; and Test 8, Tables. The hardcopy group performed significantly better than the negative-image group only, on the 25-minute narrative exercise.

For Group I (80-100 AFQT percentile scores) the hardcopy mean was significantly higher than both the positive and negative microfiche means for Test 4, Figure Identification, and significantly higher than the negative-image means for Test 6, Symbol Translation; Test 7, Graphs; and Test 11, Flow Diagrams. All of the differences reported above were significant beyond the .01 level. Again, none of the differences between means for positive versus negative-image presentations were significant for any of the AFQT groups examined.

Table IX. Scheffe Values for the Twelve Tests
(80-100 AFQT Scores)

	Test	Hardcopy versus Positive	Hardcopy versus Negative	Positive versus Negative
1.	Narrative (50)	.6974	.5138	.0213
2.	Narrative (25)	.6490	1.7392	.2457
3.	Short Paragraph	.4108	.0002	.4428
4.	Figure Identification	6.6700*	15.5770*	1.6944
5.	Length Estimation	2.4750	1.7297	.0138
6.	Symbol Translation	3.3000	8.5259*	1.0982
7.	Graphs	1.5000	7.2421*	2.2048
8.	Tables	.0963	3.7293	2.5826
9.	Mechanical Drawing	.4485	.1699	.8310
10.	Schematics	.0321	.0040	.0157
11.	Flow Diagrams	1.6144	5.7800*	1.2417
12.	Number Verification	1.2356	.6742	.1145

* Significant beyond the .10 level (4.94)

Table X. Scheffe Values for the Twelve Tests
(0-49 AFQT Scores)

	Test	Hardcopy versus Positive	Hardcopy versus Negative	Positive versus Negative
1.	Narrative (50)	4.4500	2.3800	.8308
2.	Narrative (25)	4.5800	6.0500*	.0643
3.	Short Paragraph	7.6100*	6.0200*	.0651
4.	Figure Identification	1.4240	4.2220	.1372
5.	Length Estimation	11.8300*	14.9700*	.2228
6.	Symbol Translation	4.4950	2.4290	.8302
7.	Graphs	15.0300*	17.4400*	.4496
8.	Tables	5.6378*	13.6450*	.1945
9.	Mechanical Drawing	.0167	.1021	.0127
10.	Schematics	.1381	2.4700	.6622
11.	Flow Diagrams	3762	.0219	.5083
12.	Number Verification	2.4135	2.8138	.0708

* Significant beyond the .10 level (5.40)

IV. DISCUSSION AND CONCLUSIONS

While this study was conceived as a replication of the Baldwin and Bailey investigation, it was executed in a manner consistent with achievement of the goals of the larger program: Development of Microform Materials for Use in Technical Training. Before presenting a discussion of the study results in the context of the larger program, the replication aspects will be developed.

Replication

Baldwin and Bailey's work was an attempt to develop psychometric instruments of known reliability and content validity for measuring an individual's ability to process each of several types of information through the visual modality. These instruments were then used to determine the influence of each of three presentation modes on the individual's information processing ability.

Their development of the test instruments was only partially successful in terms of the original design. Three test instruments, evaluating the subjects' ability to read continuous prose, were validated by factor analysis. These instruments required 100 minutes of the total 162 minutes allocated for all tests. The remaining nine instruments, taken in groups of three, were designed to measure either comparative visual judgment or perceptual speed. These instruments were not validated under the Baldwin and Bailey *a priori* groupings. However, taken as a single group of nine, these instruments did have significant factor loadings for either comparative visual judgment or perceptual speed. The content and validity of these two groupings is very important in understanding the replication achieved in the present study.

Baldwin and Bailey found no significant difference in performance using the narrative materials among any of the presentation modes (Hardcopy, Positive-fiche, Negative-fiche). This result was confirmed by the present study in terms of the overall experimental results. They found no significant difference in performance across all instruments for the positive versus the negative fiche presentation. This result was also confirmed. They did, however, find significant differences between fiche and hardcopy performance for three of the nine instruments in the non-narrative grouping, all requiring precise visual discrimination. This result was confirmed, but three additional instruments also showed significant performance differences between the hardcopy and one or both of the fiche presentations.

In general then, the Baldwin and Bailey results were replicated. Their main results could be extended to a different environment, a different time, and to different equipment. However, the discovery of the additional instruments showing significant differences in performance between hardcopy and the film presentations in the present study requires elaboration.

Three general observations must be made in order to provide proper context. (1) The test instruments legislate performance differences primarily in terms of speed.* Where accuracy does enter the performance comparisons, it (accuracy) is independent of the presentation mode and is clearly dependent on content of the test instrument. (see Table V.) (2) While statistically significant differences in performance were found for several of the test instruments,

* Baldwin and Bailey also report this observation.

there is clearly no fundamental breakdown in comparative performance between fiche and hardcopy presentations. (3) The performance of a complex task, using a reader presentation, is not strictly equivalent to the performance of the same task using hardcopy only. In the test environment, the reader text presentation and the exercise answer sheet could not be spacially related in the same way as could the hardcopy text presentations and answer sheets. This split-function task alone causes some performance decrement with a film presentation. It is felt that the general trend toward poorer mean scores with the fiche presentations reflects this decrement. (See Table VII for trend details.) The importance of this observation can be seen if extremes are cited. In the narrative (Test I), the subjects supplied approximately one answer each minute for 50 minutes; in Test 6, symbol translation, the subjects supplied approximately 12 answers each minute for 10 minutes. All the tests in the second group of nine have strong split-function characteristics.

Against this background, there are two methodological differences in the studies that might explain the occurrence of significant performance differences in more of the non-narrative test exercises. First, in Baldwin and Bailey's study only three test sequences were used, with the narrative material always presented first. In the present study, the test exercises were presented in 15 different but counterbalanced sequences so that the narrative material might be found at the beginning, middle, or end of a sequence. Second, the Baldwin and Bailey study apparently tested in a sequence of presentation modes; in the present study, all modes were test concurrently. It is felt that these procedural differences alone are sufficient to explain why three more of the exercises resulted in significant performance differences.

It is plausible however, to hypothesize that this "added sensitivity" was due in part to differences in the distribution of AFQT scores in the populations sampled. An analysis (discussed below) indicated that different intelligence groups vary in their response to the method used to present technical training material. Therefore, alterations in the composition of populations with regard to intelligence groupings will affect the overall sensitivity of the test. For example, the results of the present study indicate that while subjects in the middle intelligence range are insensitive to differences in the mode of presentation, significant differences favoring hardcopy occurred in the higher intelligence group on the same three exercises which resulted in significant differences in the Baldwin and Bailey work. The inclusion of lower intelligence subjects as 20% of the population in the present study resulted in significant differences on three additional test exercises. It is hypothesized, therefore, that the distribution of AFQT scores in the population used in the Baldwin and Bailey study was skewed to omit a certain number of subjects in the 0-49 range, thus decreasing the sensitivity of their testing as compared to the present study.

In order to complete the comparative aspects of the two studies, it should be pointed out that the mean scores for the narrative and short paragraph exercises in the present study were substantially higher than those reported in the original study for the same tasks. These discrepancies are possibly due to differences in the testing procedure or environmental situation, and also might reflect differences in the composition of the sample groups. Information concerning such discrepancies, however, is insufficient to make further comment.

Appendix D presents a table of means and standard deviations for the twelve tests obtained by Baldwin and Bailey. These may be compared with Table III.

Microform Materials for Technical Training

The results of the study have direct implication for the development of microform materials to be used in training. The analysis dealing with the question of whether the mode of presentation differentially affects various intelligence groups, is particularly useful. For example, all of the significant differences found between student performance using the hardcopy presentation and student performance using a microfiche presentation occurred in the higher and lower intelligence groups. No significant differences were found on any of the test exercises between hardcopy and microfiche for the intermediate intelligence group (50-79 percentile). The lower intelligence group (0-49 percentile) was affected most strongly by the mode of presentation. Hardcopy subjects in this group performed significantly better than subjects in either the positive or negative microfiche groups on four of the twelve test exercises (Short Paragraph Narrative, Length Estimation, Tables, and Graphs) and significantly better than the negative microfiche group on the 25-minute narrative exercise. This was the only analysis which led to significant differences on any of the narrative materials. These differences suggest a possible limitation for training applications for students in lower intelligence groups using microfiche equipment. Further research, using a larger sample of subjects in each intelligence classification would be necessary to identify these limitations more adequately.

The performance of the higher intelligence group (80-100 percentile) also was affected by the mode of presentation, but less

dramatically. In this group, performance on the hardcopy presentation was significantly better than performance on either the positive or negative-image microfiche presentations on the Figure Identification exercise and better than the negative-image group only on the Symbol Translation, Graphs, and Flow Diagrams exercises. These exercises are very much dependent on the recognition of individual type characters, or precise visual discrimination.

While the recognition of this performance-intelligence link is important, one very positive result was the discovery that the intermediate intelligence group (50-79 percentile) could both read and perform tasks requiring visual judgement and perceptual speed without significant performance decrement utilizing the reader presentations.

The post-hoc analysis which operationally defined performance indicated that the test instruments could be characterized as primarily speed tests. That is, the significant differences encountered in the analyses of variance and the Scheffe tests reflect primarily the effects of the reader presentation on the speed at which the students worked. This result indicates that accuracy can be maintained using microforms and does not, in itself, preclude their use in Air Force technical training programs. It simply means that where appropriate, more time should be allowed for the completion of split-function tasks. Further research should be conducted, however, which is designed specifically to reflect machine effects on speed and accuracy as separate factors.

Finally, comment is appropriate relative to the question of image polarity. The use of the Scheffe test in the present study allowed the comparison of all pairs of means in the analysis since it has no requirement that the post-hoc comparisons be independent. As in the

previous study, no significant performance differences were found between the positive-image and the negative-image fiche presentations on any of the twelve test exercises. However, a comparison of the hardcopy means with each of the microform means indicated that on two of the test exercises (tables and graphs), the means for the hardcopy group were significantly higher than the means for the negative-image group but not significantly different from the means of the positive-image group. This finding documents a general tendency over all test instruments for the negative-image group to be less effective in the performance of the tasks than the positive-image group.

Since the factor analysis conducted by Baldwin and Bailey indicated that nine of the twelve test exercises had high loadings on the comparative visual judgement and perceptual speed factors, the above results also comment on the question of visibility versus readability discussed earlier. Given the split-function complexity added to the tasks by using any reader presentation, the above results appear to indicate that positive-image microforms are better able to meet the visibility (individual character recognition) requirements than are negative-image microforms. This type of interpretation is supported by the performance of the higher intelligence group (80-100 percentile) which had particular difficulty with the negative-image presentation (see Tables VI and IX). However, the positive-image presentation resulted in the poorest performance by subjects in the lower intelligence group (0-49 percentile). This ambiguity is perhaps partially explained by differential effects of the negative-image as a novel approach among various intelligence groups or by differences in previous experience. Further research should be conducted to clarify the issue of image polarity by examining the role of image degradations

(or visibility differences) directly as they affect student performance over various presentation modes. By testing with successively poorer visibility, the interaction between image degradation and image polarity could be documented.

REFERENCES

Baldwin, T. S. and Bailey, L. J. "Readability of Technical Training Materials Presented on Microfiche Vs. Offset Copy." Journal of Applied Psychology. Vol. 55, No. 1, 1971 (Feb), 37-41.

Ferguson, G. E. Statistical Analysis in Psychology and Education. 2nd ed. New York: McGraw-Hill Book Co., 1966.

Kit of Reference Tests for Cognitive Factors. Princeton, New Jersey: Educational Testing Service, 1963.

Klare, G. The Measurement of Readability. Iowa State University Press, 1963.

Kottenstette, J. P. Study of the Characteristics of Ultramicrofiche and Its Application to Colleges and Universities. Denver: University of Denver, 1969 (Interim Report).

Myers, J. L. Fundamentals of Experimental Design. Boston: Allyn and Bacon, Inc., 1966.

Scheffe, H. The Analysis of Variance. New York: John Wiley and Sons, Inc., 1959.

Uhlaner, J. E. "Development of Armed Forces Qualification Test and predecessor army screen tests, 1946-1950," PRB Report 976, 1952.

Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill Book Co., 1962

APPENDIX A

TEST INSTRUCTIONS AND SAMPLE ITEMS

The following pages consist of reading matter which you will be asked to read for both speed and accuracy. On each page three words have been eliminated and replaced by blanks. Your task will be to read the material and supply the missing words. Read the following example and identify the missing word.

> Survivors must know how to exploit to their advantage the meat of game and fish, how to prepare plant food for eating and above all how to accomplish this with the least effort and physical exertion. Many men have died from starvation because they have failed to take full advantage of a game carcass or the plant ____1____ available.

In the practice example the word "food" was eliminated. You would therefore write "food" on the separate answer sheet in the space corresponding to the number of the blank.

When you are told to begin, turn the page immediately and begin to work. You will be allowed 50 minutes for this exercise. You are not expected to complete the test. Do not spend too much time on any one word, but do not hurry. Attempt to read at a steady pace.

No words are eliminated on the first page to allow you to familiarize yourself with the material. Beginning on page two, record on the separate answer sheet the missing word corresponding to the number in the blank. Please do not mark on this test booklet.

TESTS 1 AND 2. NARRATIVE

The following exercises consist of short paragraphs in which one word has been changed. Your task will be to read the paragraphs and identify the word which is not consistent with the meaning of the paragraph. Read the following practice example and identify the inappropriate word.

> Survivors must know how to exploit to their advantage the meat of game and fish, how to prepare plant food for eating and above all how to accomplish this with the least effort and physical exertion. Many men have died from drowning because they failed to take full advantage of a game carcass or the plant food available.

In the practice example the word "drowning" does not agree with the context of the rest of the paragraph. You would therefore write the word "drowning" on the separate answer sheet to indicate the incorrect word.

Now read the following additional practice examples. This time write the incorrect word on the separate answer sheet.

1. The sea of the Arctic Basin and the shores adjoining it have few fish or shellfish useful for survival purposes. The inland lakes and rivers of the surrounding coastal tundra, however, generally have plenty of animals which are easy to catch during the warmer season.

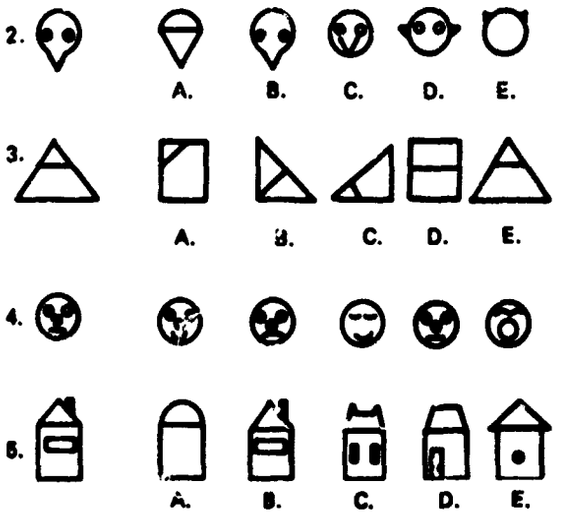
2. If mussels are the only available food, select only those in deep inlets far from the coast. Remove the dark intestinal gland after eating.

TEST 3, NARRATIVE - SHORT PARAGRAPH

This is a test of your ability to match a given object. At the left of each row is an object. To the right are five test objects, one of which matches the object at the left. Look at the example below:

The third test object (C) is the correct response, because it is the same as the object at the left.

Now practice on the problems below. Circle the letter on the separate answer sheet for the object that matches the one at the left. Make no marks on the test booklet.



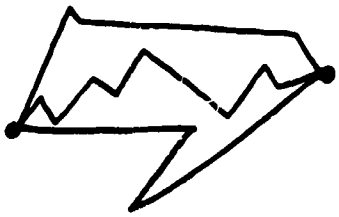
The correct responses for the practice exercises above should be: 2 = B; 3 = E; 4 = D; and 5 = B.

When you are told to begin, turn the page and immediately begin to work. You will have ___ minutes for each of the two exercises for this part. Work as quickly as you can without sacrificing accuracy. Each exercise has two pages. If you finish Exercise 1, STOP. Please do not go to Exercise 2 until you are asked to do so.

TEST 4. FIGURE IDENTIFICATION

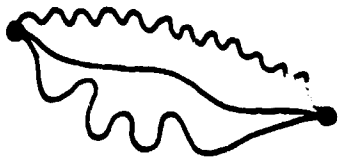
In this test you are to examine three lines or roads that connect pairs of points (•) and select the line that is the shortest. The top or high road is A; the middle road is B; and the bottom or low road is C. Look at the two sample problems below. On the separate answer sheet mark the shortest road by circling the correct letter under each problem. Make no marks on the test booklet.

I.



A B C

II.



A B C

In the examples above, A is the answer to problem I; and B is the answer to problem II.

When you are told to begin, turn the page and immediately begin to work. You will have ___ minutes for each of the two exercises for this part. Work as quickly as you can without sacrificing accuracy. Each exercise has one page. If you finish Exercise 1, STOP. Please do not go to Exercise 2 until you are asked to do so.

TEST 5, LENGTH ESTIMATION

This is a test of your ability to translate symbols into alphabetic characters. Following is a partial list of typewriter characters (symbols) that correspond to letters of the alphabet.

SYMBOL	LETTER	SYMBOL	LETTER
]	A	%	F
[	B	†	G
@	C	&	H
!	D	*	I
$	E	)	J

In this exercise you will be required to write in the blanks the alphabetic character that corresponds to the given symbol. The answer will consist of jumbled letters rather than actual words. Refer to the above list of characters and complete the following practice examples. Record your answers on the separate answer sheet. Make no marks on the test booklet.

1. @ & *) % ____ ____ ____ ____ ____

2. !] $ † ____ ____ ____ ____

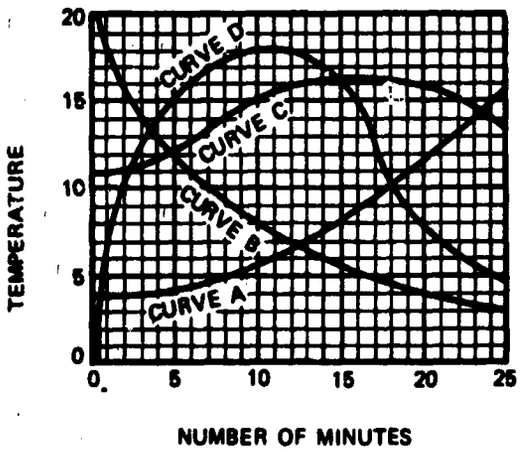
3. [% @ * ____ ____ ____ ____

The correct responses for the examples above are: 1 = CHIJF; 2 = DAEG; and 3 = BFCI.

When you are told to begin, turn the page and immediately begin to work. There is only one page to this part. You will have ___ minutes to complete the exercise. Work as quickly as you can without sacrificing accuracy. If you finish the exercise, STOP.

TEST 6. SYMBOL TRANSLATION

This is a test of your speed and accuracy in reading a graph. The graph illustrated below contains four curves — A, B, C, and D. These curves plot the relationship of temperature to time under certain experimental conditions. Minutes are marked on the lower (horizontal) edge of the graph and temperatures are shown at the left (vertical) edge. Study the example below.



In this exercise, the minutes will be given. You are to find the temperature, corresponding to the given number of minutes, for each of the four curves.

Procedure: Locate the desired number of minutes on the horizontal scale. Follow the vertical line up from that point to where it crosses Curve A. Then read the temperature from the scale at the left. Repeat the procedure for each of the other three curves. Round your answer off to the nearest whole number. Now, complete the following practice exercises. Record your answers on the separate answer sheet.

	Minutes	Curve A	Curve B	Curve C	Curve D
1.	18	____	____	____	____
2.	7	____	____	____	____

The correct responses for the exercises above are: 1. = 10, 5, 16, 10; and, 2. = 5, 10, 13, 17.

When you are told to begin, turn the page and immediately begin to work. There is only one page to this part. You will have ___ minutes to complete the exercise. Work as quickly as you can without sacrificing accuracy. If you finish the exercise, STOP.

TEST 7, GRAPHS

In this test you will be required to locate and record values found on a table. The table below indicates drill sizes and their corresponding decimal equivalents. Note that the drill sizes may be represented as letters, numbers or fractions.

SIZE	DECIMAL EQUIVALENT
1/2	0.500
Y	0.404
X	0.397
25/64	0.3906
2	0.221
130	0.185

You will be required to furnish the decimal equivalent for each of the drill sizes indicated. Your responses will be recorded on the separate answer sheet. Refer to the above table and complete the following examples.

1. 25/64 _____
2. 2 _____
3. Y _____

In the examples above the correct responses are: 1. = 0.3906; 2. = 0.221; and, 3. = 0.404.

When you are told to begin, turn the page and immediately begin to work. There is only one page to this part. You will have ___ minutes to complete the exercise. Work as quickly as you can without sacrificing accuracy. If you finish the exercise, STOP.

TEST 8, TABLES

In this test you will be required to locate lines and surfaces on three separate views of a drawing which correspond to lines and surfaces on a pictorial view of the same object. In Figure 1. below you will observe a three dimensional view (pictorial) of a box. The box is hinged on three sides so that it may be spread flat. Figure 2. is a drawing of the same box with the top raised and the side swung around to the front.

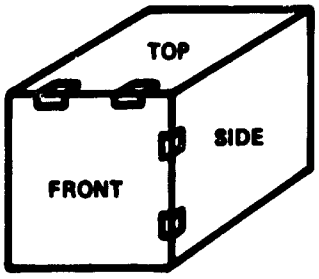


FIG. 1

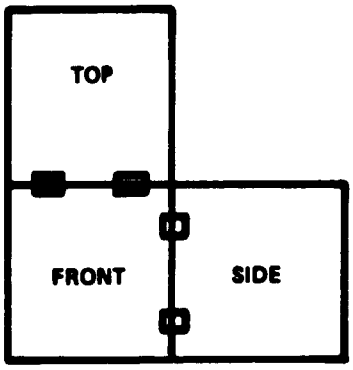


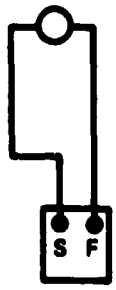
FIG. 2

The drawings which follow will utilize this same format. You will see a pictorial view of an object. Then, the object will be shown as it would appear if it were hinged and spread flat on the page. Are there any questions related to the first example?

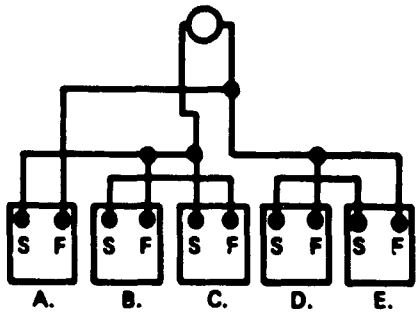
The next drawing is like the one shown above except that a part of the solid block is cut away. The same object is drawn in three views except that the "hinges" are omitted and the views are separated slightly to make it easier to visualize. Study the example on the following page to make sure the views are clearly understood. Do not, at this point, be concerned about the numbers and letters on the various views.

TEST 9, MECHANICAL DRAWING

This is a test in which you are to choose a correct path from among several choices. In the picture below is a box with dots marked S and F. S is the starting point and F is the finish. You are to follow the line from S, through the circle at the top of the picture and back to F.



In the problems in this test there will be five such boxes. Only one box will have a line from the S, through the circle, and back to the F in the same box. Dots on the lines show the only places where connections can be made between lines. If lines meet or cross where there is no dots, there is no connection between the lines. Now attempt the following example by identifying the box which has the line through the circle. Make no marks on the test booklet or answer sheet.



The first box is the one which has the line from S, through the circle, and back to F. The space lettered A would therefore have been circled on the answer sheet.

Each diagram in the test has only one box which has a line through the circle and back to the F. Some lines are wrong because they lead to a dead end. Some lines are wrong because they come back to the box without going through the circle. Some lines are wrong because they lead to other boxes that do not have lines going through the circle.

TEST 10, SCHEMATICS

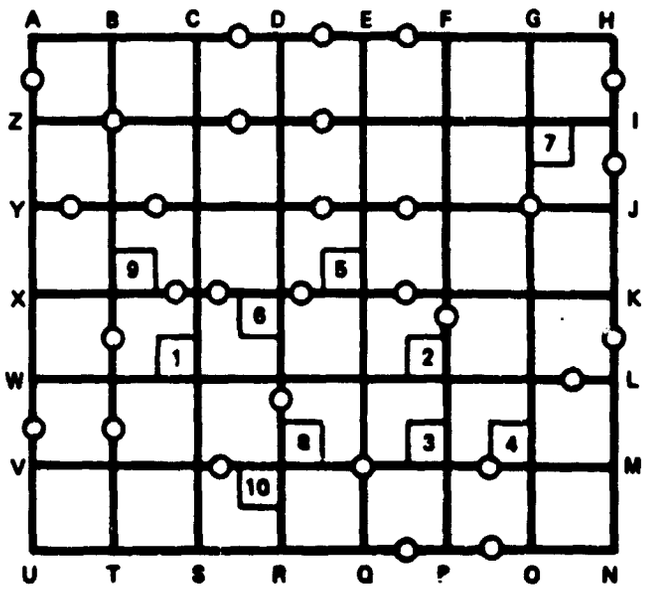
This is a test to find the shortest route between two places as quickly as possible. The drawing below is a map of a city. The dark lines are streets. The circles are road-blocks, and you cannot pass at the places where there are circles. The numbered squares are buildings. You are to find the shortest route between two lettered points. The number on the building passed is your answer.

Rules: 1. The shortest route will always pass along the side of one and only one of the numbered buildings.

2. A building is not considered as having been passed if a route passes only a corner and not a side.

3. The same numbered building may be used on more than one route.

Look at the sample map below. Practice by finding the shortest route between the various points listed at the right of the map. Your answer is to be recorded on the separate answer sheet. The first problem has been marked correctly.



The shortest route from:	Passes building:
1. A to Z	1
2. E to S	
3. P to J	
4. V to K	
5. O to F	
6. G to M	
7. D to Q	
8. F to T	

The answers to the other practice problems are as follows: 2 passes 5; 3 passes 3; 4 passes 2; 5 passes 4; 6 passes 4; 7 passes 6; and 8 passes 5.

TEST 11, FLOW DIAGRAMS

This is a test to find out how quickly you can compare two numbers and decide whether or not they are the same. If the numbers are the same, go on to the next pair, making no mark on the answer sheet. If the numbers are not the same, circle the number on the separate answer sheet corresponding to the number at the left of the incorrect pair. Now, complete the following practice examples. Make no marks on the test booklet.

1.	659 — 659	11.	7343801 — 7343801
2.	73845 — 73855	12.	18824 — 18824
3.	1624 — 1624	13.	705216831 — 795216831
4.	438 — 436	14.	971 — 971
5.	4821456 — 4821459	15.	446014721 — 446014721
6.	658331 — 656331	16.	5173869 — 5172869
7.	11653 — 11652	17.	643001717 — 6430017
8.	617439428 — 617439428	18.	518198405 — 518168045
9.	1860439 — 1860439	19.	55179 — 55097
10.	90776105 — 90716105	20.	63216067 — 63216057

The incorrect pairs in the practice examples are as follows: 2, 4, 5, 6, 7, 10, 13, 16, 17, 18, 19, and 20. These numbers should have been circled on the answer sheet.

When you are told to begin, turn the page and immediately begin to work. You will have ___ minutes for each of the two exercises for this part. Work as quickly as you can without sacrificing accuracy. Each exercise has one page. If you finish Exercise 1, STOP. Please do not go to Exercise 2 until you are asked to do so.

TEST 12, NUMBER VERIFICATION

APPENDIX B

ORDERS OF PRESENTATION

ORDER NO. 1	ORDER NO. 2	ORDER NO. 3
Introduction	Introduction	Introduction
1. Schematics (15)	1. Schematics (15)	1. Short Paragraph (25)
2. Flow Diagrams (7)	2. Flow Diagrams (7)	2. Graphs (7)
3. Number Verification (6)	3. Number Verification (6)	3. Tables (4)
10-minute break	10-minute break	4. Mechanical Drawing (5)
4. Narrative (50)	4. Graphs (7)	20-minute break
5. Narrative (25)	5. Tables (4)	5. Schematics (15)
20-minute break	6. Mechanical Drawing (5)	6. Flow Diagrams (7)
6. Graphs (7)	7. Short Paragraph (25)	7. Number Verification (6)
7. Tables (4)	10-minute break	10-minute break
8. Mechanical Drawing (5)	8. Narrative (25)	8. Figure Identification (5)
9. Figure Identification (5)	9. Narrative (50)	9. Length Estimation (3)
10. Length Estimation (3)	20-minute break	10. Symbol Translation (10)
11. Symbol Translation (10)	10. Figure Identification (5)	10-minute break
10-minute break	11. Length Estimation (3)	11. Narrative (50)
12. Short Paragraph (25)	12. Symbol Translation (10)	12. Narrative (25)

ORDERS OF PRESENTATION (Continued)

ORDER NO. 4	ORDER NO. 5	ORDER NO. 6
Introduction	Introduction	Introduction
1. Short Paragraph (25)	1. Narrative (50)	1. Graphs (7)
10-minute break	2. Narrative (25)	2. Tables (4)
2. Schematics (15)	20-minute break	3. Mechanical Drawing (5)
3. Flow Diagrams (7)	3. Graphs (7)	4. Short Paragraph (25)
4. Number Verification (6)	4. Tables (4)	10-minute break
10-minute break	5. Mechanical Drawing (5)	5. Schematics (15)
5. Narrative (50)	6. Schematics (15)	6. Flow Diagrams (7)
6. Narrative (25)	10-minute break	7. Number Verification (6)
20-minute break	7. Flow Diagrams (7)	10-minute break
7. Graphs (7)	8. Number Verification (6)	8. Narrative (50)
8. Tables (4)	9. Figure Identification (5)	9. Narrative (25)
9. Mechanical Drawing (5)	10. Length Estimation (3)	20-minute break
10. Figure Identification (5)	11. Symbol Translation (10)	10. Figure Identification (5)
11. Length Estimation (3)	10-minute break	11. Length Estimation (3)
12. Symbol Translation (10)	12. Short Paragraph (25)	12. Symbol Translation (10)

ORDERS OF PRESENTATION (Continued)

ORDER NO. 7	ORDER NO. 8	ORDER NO. 9
Introduction	Introduction	Introduction
1. Figure Identification (5)	1. Short Paragraph (25)	1. Narrative (50)
2. Length Estimation (3)	10-minute break	2. Narrative (25)
3. Symbol Translation (10)	2. Schematics (15)	20-minute break
4. Short Paragraph (25)	3. Flow Diagrams (7)	3. Figure Identification (5)
10-minute break	4. Number Verification (6)	4. Length Estimation (3)
5. Schematics (15)	10-minute break	5. Symbol Translation (10)
6. Flow Diagrams (7)	5. Narrative (25)	6. Schematics (15)
7. Number Verification (6)	6. Narrative (50)	10-minute break
10-minute break	20-minute break	7. Flow Diagrams (7)
8. Narrative (50)	7. Graphs (7)	8. Number Verification (6)
9. Narrative (25)	8. Tables (4)	9. Short Paragraph (25)
20-minute break	9. Mechanical Drawing (5)	10-minute break
10. Graphs (7)	10. Figure Identification (5)	10. Graphs (7)
11. Tables (4)	11. Length Estimation (3)	11. Tables (4)
12. Mechanical Drawing (5)	12. Symbol Translation (10)	12. Mechanical Drawing (5)

ORDERS OF PRESENTATION (Continued)

ORDER NO. 10	ORDER NO. 11	ORDER NO. 12
Introduction	Introduction	Introduction
1. Short Paragraph (25)	1. Graphs (7)	1. Schematics (15)
2. Graphs (7)	2. Tables (4)	2. Flow Diagrams (7)
3. Tables (4)	3. Mechanical Drawing (5)	3. Number Verification (6)
4. Mechanical Drawing (5)	4. Figure Identification (5)	10-minute break
10-minute break	5. Length Estimation (3)	4. Short Paragraph (25)
5. Narrative (50)	6. Symbol Translation (10)	10-minute break
6. Narrative (25)	10-minute break	5. Figure Identification (5)
20-minute break	7. Narrative (50)	6. Length Estimation (3)
7. Schematics (15)	8. Narrative (25)	7. Symbol Translation (10)
8. Flow Diagrams (7)	20-minute break	8. Graphs (7)
9. Number Verification (6)	9. Short Paragraph (25)	9. Tables (4)
10-minute break	10-minute break	10. Mechanical Drawing (5)
10. Figure Identification (5)	10. Schematics (15)	20-minute break
11. Length Estimation (3)	11. Flow Diagrams (7)	11. Narrative (50)
12. Symbol Translation (10)	12. Number Verification (6)	12. Narrative (25)

ORDERS OF PRESENTATION (Continued)

ORDER NO. 13	ORDER NO. 14	ORDER NO. 15
Introduction	Introduction	Introduction
1. Narrative (50)	1. Graphs (7)	1. Schematics (15)
2. Narrative (25)	2. Tables (4)	2. Flow Diagrams (7)
20-minute break	3. Mechanical Drawing (5)	3. Number Verification (6)
3. Short Paragraph (25)	4. Figure Identification (5)	10-minute break
10-minute break	5. Length Estimation (3)	4. Narrative (50)
4. Figure Identification (5)	6. Symbol Translation (10)	5. Narrative (25)
5. Length Estimation (3)	10-minute break	20-minute break
6. Symbol Translation (10)	7. Short Paragraph (25)	6. Short Paragraph (25)
7. Schematics (15)	10-minute break	10-minute break
10-minute break	8. Narrative (25)	7. Graphs (7)
8. Flow Diagrams (7)	9. Narrative (50)	8. Tables (4)
9. Number Verification (6)	20-minute break	9. Mechanical Drawing (5)
10. Graphs (7)	10. Schematics (15)	10. Figure Identification (5)
11. Tables (4)	11. Flow Diagrams (7)	11. Length Estimation (3)
12. Mechanical Drawing (5)	12. Number Verification (6)	12. Symbol Translation (10)

APPENDIX C

COMPUTING THE SCHEFFE METHOD OF POST-HOC COMPARISONS

The following method of making selected a posteriori and complete sets of comparisons among experimental group means was developed by Scheffe (5). A simple means of applying this method is suggested by Furguson (4) according to the following procedure.

Step 1. Calculate F ratios using the following formula:

$$F = t^2 = \frac{(\bar{X}_1 - \bar{X}_2)^2}{Sw^2/n_1 + Sw^2/n_2} = \frac{(\bar{X}_1 - \bar{X}_2)^2}{Sw^2(n_1 + n_2)/n_1 n_2}$$

where Sw^2 equals the within-group variance, and n equals the sample size.

Step 2. Consult a table of F and obtain the value of F required for significance at the desired level for $df_1 = k-1$ and $df_2 = N-k$.

Step 3. Calculate a quantity F', which is k-1 times the F required for significance at the desired level; that is, $F' = (k-1)F$.

Step 4. Compare the values F and F'. For any difference to be significant at the desired level, F must be greater than or equal to F'.

Appendix D. Test Means and Standard Deviations in the Baldwin and Bailey Study

Test No.		Offset Copy N=45		Positive N=43		Negative N=45	
		$\bar{X}$	S.D.	$\bar{X}$	S.D.	$\bar{X}$	S.D.
* 1.	Narrative (50 min.)	37	15.2	38	19.1	40	20.7
* 2.	Narrative (25 min.)	19	7.5	19	8.7	20	11.2
* 3.	Narrative-Short Paragraph	27	11.8	27	11.7	30	13.3
4.	Figure Identification	61	12.5	54	9.0	52	8.7
5.	Length Estimation	30	6.0	28	8.8	27	6.2
6.	Symbol Translation	141	28.4	114	28.3	117	30.7
7.	Graphs	34	10.8	28	9.6	29	9.9
8.	Tables	17	5.7	16	5.7	15	6.0
9.	Mechanical Drawing	27	16.4	33	18.3	36	20.4
10.	Schematics	10	5.7	11	5.7	12	6.4
11.	Flow Diagrams	19	5.5	18	4.1	20	5.1
12.	Number Verification	30	6.9	29	6.9	29	7.3

* Tests in which mean scores were substantially lower in the Baldwin and Bailey study than in the present study (see Table III).