RADC-TDR-63-421

RAPID IDENTIFICATION AND INTERPRETATION TECHNIQUES

TECHNICAL DOCUMENTARY REPORT NO. RADC-TDR-63-421

November 1963

Information Processing Laboratory
Rome Air Development Center
Research and Technology Division
Air Force Systems Command
Griffiss Air Force Base, New York

Project No.6244, Task No.624401

(Prepared under Contract No. AF30(602)-2806 by Richard L. Krumm and Alfred J. Farina, Jr, Radio Corporation of America, Data Systems Center, Bethesda, Maryland)

FOREWORD

Appreciation is expressed to the United States Naval Photographic Interpretation Center, Suitland, Maryland, and the United States Army Personnel Research Office, Washington, D. C., for providing test subjects for purposes of this study. The interest and cooperation of their skilled personnel contributed greatly to the study results.

Appreciation is also due the expert intelligence analysts who developed the scoring keys for the imagery. Mr. Roger Graves of RCA Data Systems Center, Mr. Bruce Herrington of Weiser Associates, and Mr. G. Bigelow and Mr. A. Schwartz of the United States Army Personnel Research Office provided invaluable assistance in this regard.

Finally, the substantial contributions of the Project Monitor, Mr. Angelo Zieno of the Rome Air Development Center, are acknowledged. Mr. Zieno and Dr. Shelton MacLeod, also of RADC, provided valuable suggestions at critical points in the research effort. In addition, Mr. Zieno accomplished the preliminary screening of the rolls of aerial imagery which were ultimately searched for project training materials. His efforts in securing, from a variety of governmental sources, imagery containing a wide variety of military targets contributed materially to the success of the project.

AIR FORCE, Cato Show Prtg. Co. 12/18/63 - 199

Suggested Keywords: Military Science, Intelligence, Photographic Intelligence

ABSTRACT

This report describes an experimental investigation of the feasibility of utilizing a rapid presentation technique to improve the target recognition performance of photo interpreters. The test imagery involved a wide variety of military targets presented at various image scales and image quality. A four day training course administered to a group of experienced photo interpreters effected 50 percent reduction in time required to identify military targets with no loss in either accuracy or completeness.

Data analysis also indicated significant interactions among interpreters measures and such display variables as image scale, target to background contrast, and image complexity.

This report will be of particular interest to persons concerned with photo interpretation research and training.

PUBLICATION REVIEW

This report has been reviewed and is approved. For further technical information on this project, contact Mr. Angelo S. Zieno, RAWIC, Ext. 26116.

Approved: Angelo S. Zieno
ANGELO S. ZIENO
Project Engineer
Directorate of Intelligence & Electronic Warfare

Approved: Frank J. Tomaini
FRANK J. TOMAINI
Chief, Information Processing Laboratory
Directorate of Intelligence & Electronic Warfare

FOR THE COMMANDER: Fred I. Diamond
for IRVING J. GABELMAN
Director of Advanced Studies

SUMMARY AND CONCLUSIONS

The research study which is reported was directed toward the goals of (a) developing and evaluating a training program designed to increase the speed with which photointerpreters can perform flash identification of targets from photographic imagery, and (b) exploring certain relationships among imagery features and the several criterion measures which had been employed in assessing the rapid identification training program.

A four-day training program was prepared, based upon a series of limited-scope preliminary experiments. The training program used projected photographic imagery and was structured around four main techniques:

(a) training in simplified graphic representations of target features (target signatures) which serve as identifying elements,

(b) training in scanning,

(c) successive reduction in the time the projected image is displayed, to a minimum duration of two seconds,

(d) provision of a variable feedback schedule during training.

The imagery used during the training course involved large, medium, and small image scales and contained a variety of targets of military, industrial, rail, airfield, port/harbor, and electrical power facilities.

The program was evaluated by training a group of 15 skilled photointerpreters in rapid identification techniques, and then comparing their performance on interpretation test criterion measures with the performance achieved by an equal number of

interpreters, matched in terms of prior training and experience but who had not been administered the rapid identification training program.

The criterion tests included ninety 35mm slide images displayed in rear projection format, and a specially prepared section of nine-inch aerial roll film displayed on a conventional light table. All image scale categories were represented on both tests and all imagery consisted of positive transparencies. The subjects' responses were scored in terms of:

(a) accuracy (right responses divided by right plus wrong responses)

(b) completeness (right responses divided by total possible right responses)

(c) efficiency (right responses divided by total time imagery was viewed)

(d) effectiveness (completeness score divided by total time imagery was viewed)

In addition to the assessment of training course effectiveness in promoting rapid identification of targets, analyses were conducted of the effects upon performance of such variables as target type, number of targets per image (complexity), target-to-background contrast (conspicuity), image scale, film density, target size, and the mode of displaying the imagery.

As a result of the various analyses, a number of conclusions were possible; these are presented in detail in Section 3 in terms of their applicability to the training program or to the image and display variables studied. General conclusions regarding major aspects of the research are summarized below.

1. The Training Program

The four-day training program resulted in significantly improved performance among the Experimental group of experienced interpreters compared with a Control group matched in terms of experience and prior training. The Experimental group was seven per cent more accurate and ten per cent more complete in their identifications than were the Control group, and they provided their responses in one-half the viewing time needed by the Control group.

The Experimental group performed significantly better than the Control group in terms of Accuracy percentage for large scale imagery. The two groups performed equally for medium and small scale imagery.

The Experimental group performed significantly better than the Control group in terms of Completeness percentage for large scale and for medium scale imagery. The two groups performed equally for small scale imagery.

Certain of the effects of the rapid recognition training using a rear projection display also applied when the groups were tested under more customary viewing conditions (light table display). Accuracy and Completeness scores were equivalent for both groups but the Experimental group performed the task significantly more rapidly than did the Control group (25 per cent time reduction).

Analysis of performance during the training course indicates that the four-day training course cannot be shortened. However, since there was no indication that a "plateau" had been reached, the course might profitably be lengthened. The Accuracy and Completeness scores were not so high as to rule out further improvement.

2. Image and Display Variables

Image scale, target-to-background contrast, and image complexity were found to be significant variables in influencing completeness of interpretation. Both image scale and target contrast interacted with target type in influencing completeness of interpretation. The effects of image complexity were reduced by the training program. The Experimental group achieved substantially higher Completeness scores at all levels of image complexity but their Accuracy was impaired at the more complex imagery levels.

Target size, either absolute length of the longest dimension or the per cent of image area encompassed by the target, interacted in an irregular manner with Completeness scores for major target types. This finding should be further investigated.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

LIST OF TABLES

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES
(Continued)

LIST OF FIGURES
(Continued)

INTRODUCTION

During recent years there has been an increasing interest among both military and civilian groups in improving imagery interpretation techniques. To a certain extent this interest has been generated by the rapid technological advances in methods for collecting imagery; methods which can provide overwhelming masses of raw pictorial data within brief time periods. The interest has also received impetus from the demonstration during the Cuba crisis that pictorial data provide an irrefutable form of intelligence information.

It can be anticipated that imagery data gathering capabilities will continue to be expanded and that there will also be a continuing need for military applications of these techniques. In addition, the advent of near real-time interpretation of transmitted imagery, while not adding directly to the problem of imagery quantity, will require techniques that similarly stress rapid and accurate assessment of pictorial materials. Consequently, the necessity for improving intelligence data reduction methods is an increasingly serious one.

For pictorial intelligence data, research interests have branched in two directions. First, attention is being devoted to automatic imagery processing techniques. A considerable body of knowledge is being built up around various concepts of pattern perception, and it can be anticipated that ultimately such equipments will successfully accomplish certain types of imagery screening functions.

Second, in addition to the study of techniques for electronic processing of pictorial data, attention has centered upon various means of improving performance of the human operator. This interest can be justified on several grounds.

It is unlikely that automatic techniques can be perfected within the near future. Consequently, improving man's performance can be considered a necessary stop-gap. Furthermore, it is unlikely, even if automatic equipments are developed, that they will always be available where they are needed, or that they can be programmed to perform reliably the most detailed of the interpreter's analytic functions.

Recent military research on improving photointerpretation performance was, at the outset, characterized by "broad brush" investigations. A number of exploratory studies were conducted in an effort to discover the areas which appeared to offer the greatest research benefits. Viewing consoles, methods of projecting image displays, and team interpretation techniques, to name but a few of the various types of studies, all received attention. As measurement of interpreter performance increased, it became apparent that information was often absent which, had it been available, could have aided in interpreting study findings. Among the necessary types of information are such items as methods to equate potential test subjects in terms of prior interpretation experience, amount of training, or level of proficiency, and fundamental research concerning the manner in which characteristics of the displayed image (i.e., its "quality") influence observer performance.

Consideration of current research efforts suggests that exploratory efforts are now tending toward the more fundamental aspects of interpreter performance. Items of current research interest appear to be less concerned with the production of elaborate viewing consoles than with a genuine concern for discovering how interpreters actually accomplish their job, what factors affect their performance, and how applied training techniques might act to overcome certain of the factors which currently limit interpreter efficiency.

The present study is one in this "second generation" series. Specifically, the research had two related goals. The first was to develop and evaluate a training program designed to increase the speed with which interpreters can perform flash identification of targets from photographic imagery. The second goal was to explore certain relationships among imagery features and the several criterion measures which had been employed in assessing the rapid identification training program.

Improvements in interpreter speed, if not accompanied by losses in accuracy or completeness, can be of immediate practical benefit. Knowledge of the nature of imagery variables which adversely affect interpreter performance can have ultimate application in influencing mission-scheduling, improving the design of interpreter consoles, and modifying interpreter reporting requirements.

Although the applicability of speed information is readily apparent, the applicability of research findings concerning interpreter perceptual processes is, unfortunately, often less clear. It is for this reason that we have elected to devote a substantial portion of this paper to presentations of rationales for selecting certain independent variables for investigation, for designing and conducting the study as we did, and for the various forms of data analyses which were conducted. Implications of the research findings are treated at length. The experiment proper is of classic design and therefore needs no elaborate treatment.

The investigation involved a substantial number of variables, both independent and dependent. In order to avoid confusion among these, we have organized this report in sections corresponding to the study goals. The first section is concerned with the development and evaluation of the rapid recognition training program. The second section is concerned with an assessment of certain factors influencing interpretation, factors which must also be considered in evaluating the results of the training program. Certain implications of the research are presented in the third section of this report.

SECTION 1

DEVELOPMENT AND EVALUATION OF THE TRAINING PROGRAM

A. TRAINING COURSE CONTENT

A primary concern in the present study was that of assuring that the training materials be representative of those of operational interest. This consideration is a basic prerequisite to direct application of study results. Consequently, a substantial portion of the study effort was devoted to the tedious task of screening quantities of aerial imagery in order to select as subject materials (a) a broad range of targets of military importance, (b) targets at various image scales, and (c) imagery which varied in "complexity," i.e., contained from one to five or more different target types in a single frame.

The nature of the selected imagery established the skeletal structure about which the program was designed. The targets selected were functionally grouped, and this grouping served as the basis for the development of target "signatures," for structuring of subject response sheets, and for the formulation of scoring techniques. The processing of the imagery influenced the selection of display modes which, in turn, led to the addition of a second means of evaluating the training program.

1. The Imagery. During the initial phase of the study, more than 75 rolls of 9x9-inch aerial film in both positive and negative transparency format were reviewed. In addition, several rolls of five-inch positive transparencies and several hundred cut films were studied. The initial screening was accomplished by a highly skilled interpreter with more than 20 years experience in the photo intelligence field.

Selected frames were photographically reduced to 35mm positive transparency format. This initial screening resulted in more than 600 glass-mounted slides. In preparing the slides, care was taken to ensure that targets of interest were not necessarily centered; the entire width of the selected frame was reduced to 35mm in order to preserve the image scale relationships.

The image pool was again reviewed. Each slide was individually studied in order to identify each target which it might contain. In this establishment of "ground truth," use was made of magnification equipment, interpretation keys, other imagery of the same area, and additional experienced photo-interpreter consultants.

The slides were sorted in terms of approximate image scale values and in terms of the types of targets they contained, preparatory to developing the training and test materials. The test imagery was selected first, 30 slides at "large," "medium," and "small" scale values.* These 90 test images represented, at each scale category, the majority of targets which were contained in the initial pool.

A similar procedure was followed in culling the training imagery. The intent was to organize slide cartridges to

* The reader is cautioned to observe that these terms apply only to the imagery used in the present study and do not correspond to the values commonly employed. The present study sorted in terms of Large (1:5000 and below); Medium (1:5100 to 1:15000); and Small (1:15100 and above). However, the image scale values were calculated without knowledge of the terrain height and hence are only approximations. More importantly, each nine-inch image was projected at a 20-inch size which more than halved the original scale value. Later analyses of the relationships of scale to criterion measures are legitimate, but critical scale values cannot be expected to apply directly to other research investigations.

progress from "easy" large scale, single target slides to "difficult" small scale, multiple target slides while maintaining an equitable balance among target types at each scale and assuring that all targets contained in the test imagery were represented in the training imagery. In addition, attention was given to density. Since a rear projection display and flash presentation was to be used, it was felt that in the interests of subject comfort the displays should be fairly uniform in terms of density. These culling operations reduced the number of slides which were actually employed during the study to slightly more than 300 images containing a total of 660 targets. A distribution of major target types by scale category for these slides is presented in Table 1-1.

TABLE 1-1.

DISTRIBUTION OF TARGET TYPES BY SCALE CATEGORY FOR THE SLIDE IMAGERY

	Large Scale		Medium Scale		Small Scale		Total	
	N	%	N	%	N	%	N	%
Military	35	5.3	27	4.1	43	6.5	105	15.9
Industry	35	5.3	86	13.0	94	14.2	215	32.6
Airfield	12	1.8	21	3.2	46	7.0	79	12.0
Railroad	28	4.2	67	10.1	79	12.0	174	26.4
Port/Harbor	6	1.0	11	1.7	49	7.4	66	10.0
Power	4	0.6	8	1.2	9	1.4	21	3.2
TOTAL	120	18.2	220	33.3	320	48.5	660	100.1

Although the employment of slides as training and testing imagery offered substantial advantages with respect to ease of programming the display sequences, timing of the exposures, and simplicity of the equipment, it also raised questions concerning the applicability of study findings. Rear projection is not yet a common display mode for image interpreters. If the training program were successful in reducing the required viewing time, would the results still be applicable when the trained subjects reverted to their accustomed (light table) viewing mode?

To provide data concerning such possible transfer effects, a second set of test materials was prepared. This consisted of 17 three-frame sequences from various 9x9-inch negative transparency aerial roll film which were spliced and converted to duplicate positives. The sequences were roughly ordered in terms of scale and, generally, were representative of the major target types included in the slide imagery. Table 1-2 presents information concerning the distribution of major target types for the three image-scale categories for both criterion tests.

2. The Targets. Following the establishment of the listing of targets contained in the imagery, the targets were grouped in terms of gross functional categories, e.g., Industrial, Rail Transportation, Military, etc. Then, within each functional grouping, four additional "levels of specificity" were identified. Considering the family grouping as the first level (e.g., Industrial), the second level defines the general type, the third level defines the specific type, the fourth level concerns the level of activity and the fifth level concerns threat evaluation and vulnerability. It appears that the first three levels require primarily an emphasis upon perceptual and cognitive skills. Fourth and fifth level functions appear

TABLE 1-2.

DISTRIBUTION OF TARGET TYPES BY SCALE CATEGORY FOR THE CRITERION TESTS

(Light Table Test)

	Large Scale		Medium Scale		Small Scale		Total	
	N	%	N	%	N	%	N	%
Military	1	1.2	7	8.4	5	6.0	13	15.7
Industry	3	3.6	12	14.5	13	15.7	28	33.7
Airfields	-	-	3	3.6	3	3.6	6	7.2
Railroads	3	3.6	8	9.6	12	14.5	23	27.7
Port/Harbor	-	-	3	3.6	4	4.8	7	8.4
Power	1	1.2	2	2.4	3	3.6	6	7.2
TOTAL	8	9.6	35	42.2	40	48.2	83	99.9

(Projected Display Test)

	Large Scale		Medium Scale		Small Scale		Total	
	N	%	N	%	N	%	N	%
Military	19	8.8	6	2.8	10	4.7	35	16.3
Industry	15	7.0	37	17.2	22	10.2	74	34.4
Airfields	4	1.9	3	1.4	10	4.7	17	7.9
Railroads	14	6.5	27	12.6	23	10.7	64	29.8
Port/Harbor	1	0.5	4	1.9	16	7.4	21	9.8
Power	-	-	1	0.5	3	1.4	4	1.8
TOTAL	53	24.7	78	36.4	84	39.1	215	100

largely during detailed interpretation tasks. The first three levels are involved in the preparation of Flash and Immediate reports. For this reason, the present study did not progress beyond level three, and study emphasis was placed upon the shape or configurational characteristics which were believed to be helpful in arriving at identifications at each of the first three levels (see Table 1-3 for a detailed listing of the types of targets used).

One of the major benefits to be derived from a hierarchical structuring such as this is that it provides valuable information concerning the level of detail that can be achieved in reporting as a function of the amount of time spent in viewing the image. If all that is asked, for example, is that an observer detect airfields, few errors will occur even among non-interpreters. The distinctions required to differentiate military and civilian airfields are, however, not always readily apparent to an untrained observer. Thus, a "ceiling" is provided to the test which is sufficiently high to permit differences among subjects to appear. On the other hand, if an observer is unable to discern the identifying elements which permit interpretation at a detailed level, he can report the detected target at the next higher abstraction level. In a test situation he receives at least partial credit; in an operational setting his efficiency may be similarly improved.

There is an additional benefit to be derived from the category system beyond that of convenience of testing. Each of the target types, regardless of functional category level, can be graphically portrayed by geometric representations which bear direct relationships to significant form features of the target. These essential identifying elements, or target "signatures," are the features which are necessary and sufficient to permit

TABLE 1-3.

TARGET NAMES AND ABBREVIATIONS

I. MILITARY (MIL)
   1. Electronics (ELEC)
      a. Radar Site (RADAR)
      b. Range Station (RS)
      c. Antennas (ANT)
   2. Missile Sites (MISS)
      a. Nike (NIKE)
      b. ICBM (ICBM)
      c. IRBM (IRBM)
   3. Explosives (EXP)
      a. Ammunition Storage (AMMO)
      b. Explosive Chemicals Storage (ECS)
   4. Military Base (MB)
   5. Military Supply/Storage Depot (MS)

II. INDUSTRY (IND)
   1. Processing (PROC)
      a. Steel Plant (STEEL)
      b. Chemical Plant (CHEM)
      c. Refinery (REF)
   2. Fabrication (FAB)
      a. Light Fabrication (LF)
      b. Heavy Fabrication (HF)
   3. Petroleum, Oil, Lubricants Tanks (POL)

III. RAILROAD (RR)
   1. Marshalling Yard (MY)
   2. Roundhouse (RH)
   3. Trans-shipment Yard (TY)

IV. AIRFIELDS (AF)
   1. Military Airfield (MAF)
   2. Civilian Airfield (CAF)
   3. Private Airfield (PAF)
   4. Plane Manufacturing (PM)

V. PORT/HARBOR (P/H)
   1. Civilian Port/Harbor (CP/H)
   2. Military Port/Harbor (MP/H)
   3. Shipyard (SY)
      a. Civilian Shipyard (CSY)
      b. Military Shipyard (MSY)

VI. POWER (PWR)
   1. Thermo-Electric Plant (TEP)
   2. Hydro-Electric Plant (HEP)
   3. Atomic Power Plant (APP)

the unequivocal assignment of an identifying name to a viewed image. These graphic representations formed a central part of the training program for this study. Some explanation of their derivation would appear to be in order.

The concept of target signatures is not new. The identifying characteristics of target types are written in all interpretation keys, frequently accompanied by photographs and drawings showing the way in which these elements appear. There are two primary reasons why this type of signature information may be unsuited for rapid identification: (a) the keys are largely designed for detailed interpretation so that the data presented are far more extensive than necessary for target identification, and (b) there is no evidence to indicate which target features are "important," in the sense that they are the necessary and sufficient ones.

The information contained in interpretation keys is largely pertinent to the analyses of function which are performed during detailed interpretation efforts, e.g., work-flow routes, building dimensions, output capacity, etc. Any wholesale conversion of this information into "target signatures" would be impractical because of the substantial difference in time requirements of the identification and interpretation tasks. Since, in a rapid identification effort, the subject would not have unlimited time to search the imagery, it would be uneconomical to require him either to memorize or consult lengthy lists of target characteristics in order to identify a target. It appeared that the development of "rapid" signatures, (i.e., a condensed listing of salient elements) would be desirable. Accomplishment of this step required a means of isolating from the wealth of detailed target information those points which were specific to rapid identification efforts. Such a step may be accomplished "logically" by abstracting from the lists of characteristics presented in interpretation keys those more prominent (hence, perhaps more readily detectable) features, or it may be approached empirically. The latter alternative course was emphasized.

The procedure involved having subjects record (draw) what they saw during the rapid presentation of imagery containing targets of interest. These drawings were then analyzed to select features which were both (a) noted by a majority of the subjects, and (b) identified as being valid recognition signs in existing PI keys. To avoid a mere parroting of what had previously been learned, it was important to use a group of subjects untrained in interpretation techniques.

Accordingly, a small study was conducted using four subjects and 36 slides representing multiple examples of selected target categories. Each of the subjects was tested separately. Each slide was presented twice, first at 0.1-second and then at a 2-second exposure duration. The subjects graphically recorded their perceptions after each exposure.

The results were illuminating in the sense of indicating that certain features which may be assumed to be essential in identifying a particular target type may not even be seen by a subject under brief display conditions. For example, from data contained in interpretation keys it may be assumed that a thermal electric power plant has at least five major identifying elements: coal pile, power station, conveyor system, stacks, and transformer yard. However, at the 0.1-second and 2-second exposures only two of these features were consistently recorded by the subjects: the coal pile and the power station. A similar "savings in elements" was found for fabrication plants, steel mills, certain types of military installations, military airfields, and ammunition storage areas. On the other hand, confirmation of "logically" derived signatures occurred for missile sites, POL, aircraft manufacturing plants, and certain rail transportation facilities.

The potential contribution of graphic signatures was explored in another small scale study. Three groups comprised of four non-interpreters each, were differentially trained. Four subjects received training* in identifying targets using slides selected from the test imagery pool. Four subjects were trained to identify graphic signatures of the same target families using only "synthetic" imagery (i.e., slides prepared by photographing cardboard cutout representations of target signatures). The third group of four subjects received no special training other than being provided with the list of targets they were to identify.

All subjects were then tested using actual imagery. The first and third group performances provided the boundaries for assessing performance of the group trained solely with graphic displays. It was found that at each image scale category the group trained on synthetic imagery, performed "better" than did the untrained group, though not so well as the group trained on actual imagery.** This transfer effect was sufficiently great to warrant the conclusion that the graphic signatures appeared to have some validity.

---

* This "training" involved displaying each target type, giving its identifying name, describing the nature of the identifying characteristics, and then providing practice (with feedback) for a series of slides containing the targets of interest.

** For example, indexes of performance for a 2-second exposure test were 40, 35, and 15, respectively, for the "real," "synthetic," and "untrained" groups.

3. Training Techniques. The training program was structured around four main techniques: the development of graphic target signatures (discussed above); scanning training; a systematic reduction in exposure duration as training progressed; and a variable reinforcement (feedback) schedule.

a. Signature Training. The method of introducing signature training into the program was essentially as follows, for each of the target types:

- The signature of the target was shown as a drawing which contained the major visual elements as determined from the above-mentioned studies
- Each element in the graphic signature was identified and discussed
- The signature elements which would "survive" under rapid display conditions were pointed out
- Associational features which might aid in identification but which might not necessarily be present in all imagery were also discussed

b. Scanning Training. In studies of visual search, a common observation is that free search scanning habits are frequently inefficient. Enoch (1960), for example, has shown that when observers are required to locate a target embedded in a complex background, scanning occurs in two stages. There is first a typical orientation search, in which the eye makes a brief excursion of the total display, followed by a second, more specific stage during which fixations are generally concentrated within a limited portion of the display. The typical observer fixates in the center of the display in this latter stage, disregarding the peripheral areas. Enoch noted that "this tendency has been found to be independent of the size of the display, the quality and content of the displayed image, and the time allowed for viewing.." Because the targets dealt with by photointerpreters are typically small, requiring foveal viewing, it would appear that "natural" scanning habits must be broken and the observer trained to achieve the greatest accuracy within a minimum exposure time.

The scanning training was administered in two sections. First, the subjects were familiarized with the typical scanning errors and were given advice on how to avoid committing them. Typical errors were cited as:

(a) Fixating on the center of the display, leaving the edges unscanned.

(b) Relaxing the search when one or two targets are identified.

(c) Moving the head and/or the entire body instead of relying mainly on eye movements.

(d) Continuing to view a detected target beyond the time required for making an identification.

In overcoming such errors, the subjects were advised to:

(a) Devote an equal amount of initial scanning time to the edges and center of the display.

(b) Assume that each slide contains six targets and keep scanning until time runs out.

(c) Move the eyes rather than the body or head.

(d) Not perform a detailed analysis - "identify and move on."

The second section of the scanning training portion consisted of practice using synthetic imagery. The target objects (nature unspecified) in a single slide varied in terms of number, size, and location. From one to five target objects could be displayed and each object could occupy one of five positions (the four quadrants and a center spot). Within a training series, the slides began with "large scale" single target imagery in which the targets were centrally located, and progressed along the dimensions of object number, size, and location so that final slides could each contain several very small target objects at various display locations. The subject's task was to record the number and location of the target objects by marking a prepared record form. The initial exposure duration was set at 1-second and a second presentation was made at 0.1-second. The intent of this type of practice was to establish a readiness on the part of the subjects to expect targets to appear anywhere in the display.

In addition to the above specific inputs to scanning training, the training in target signatures may also be considered as a potential contributor to efficient scanning behavior. The signature information reduced the number of elements, or critical items, to be detected in order for an identification to be made, and thus (presumably) allowed more of the display to be scanned in a specified time period.

c. Exposure Duration Reduction. This technique involved the repeated showing of selected slides at regularly decreasing exposure times. For the study's purpose, 8-seconds was arbitrarily set as the maximum time permissible during training and the ultimate goal was established as achieving a 2-second duration with no loss in accuracy or completeness of target identification. The specific exposures used are presented in the chart that follows.

	Presentation Sequence			
Imagery	1	2	3	4
Part I: 4 cartridges of slides	8 sec.	6 sec.	4 sec.	2 sec.
Part II: 3 cartridges of slides	4 sec.	3 sec.	2 sec.	2 sec.

As indicated in the chart, Part I training employed four cartridges of slides to which the subjects were first exposed at an 8-second duration per slide. The subjects recorded their identifications in a booklet after each exposure using the target abbreviations to conserve training time. Each slide was immediately succeeded by a "synthetic" representation of the image, providing feedback on the adequacy of the subject's response. Then the original slide was reprojected briefly.

This same series of four cartridges was then presented at 6-, 4-, and 2-second exposure durations. Each slide was thus viewed for a total of 20 seconds. Any performance improvement on successive trials within Part I represented, at best, a combination of recall of previously detected targets as well as an increased efficiency.

Part II training was essentially the same as Part I training except that three cartridges of "new" slides were introduced, with the exposure duration set at 4-seconds for the first exposure of three cartridges, 3-seconds for the second exposure, and 2-seconds for the third and fourth exposures.

Two "tests" were administered, one after Part I and one after Part II training. Each of these measures consisted of a cartridge of "new" imagery and served to check whether improvements in subject performance were being achieved. The tests may be considered as comprising part of the training program since the test images were later reprojected at a 2-second exposure and feedback was verbally provided following each slide display.

d. Feedback. The type of reinforcement employed during Part I and Part II training consisted of the presentation of a synthetic image immediately after the display of a real image, and then a redisplay of the real image. The synthetic image contained signature drawings, the abbreviated names of the targets which were present in the slide, and the locations of the targets in the display. One advantage of using synthetic slides was that the per cent of feedback could be systematically and automatically reduced. The percentage of slides for which the subject received feedback ranged from 100% for each of the slides in the four cartridges of Part I to 50%, 25%, and 0%, respectively, for the three cartridges of Part II.

B. DEVELOPMENT OF THE CRITERION MEASURES

Even a cursory consideration of the role of the military interpreter indicates that several standards exist against which his performance must be assessed. Foremost among these is speed, both because of certain military requirements and because of the large mass of incoming pictorial data.

In addition, it is critically important that the interpreter be accurate in his identifications. Accuracy in this sense refers to assigning the correct name to a perceived object, whether it be a decoy target which resembles an aircraft or whether it be an actual camouflaged equipment.

Thirdly, there is a requirement for completeness. Whereas accuracy is concerned with reducing errors of commission, completeness is concerned with reducing errors of omission. However, in addition to the need for reporting all targets, completeness also implies the reporting or identification of targets to the most specific descriptive terms possible under the existing viewing conditions.

Finally, either accuracy or completeness can be considered in terms of speed. That is, the interpreter who provides more correct responses per unit time can be considered to be operating more effectively than an interpreter who (a) may have as many correct responses but takes longer to produce them or (b) one who works as rapidly but provides fewer correct responses. The term "effectiveness" is used in the present report only for the measure of completeness per unit time.

In addition to these sorts of considerations, there was at issue in the present study a potential problem concerning the display format. The bulk of the developed training program dealt with projected images, a display format with which few interpreters are familiar. Will performance criteria developed for such equipment apply meaningfully to interpreter tasks performed under different display conditions? Obviously, this factor must be explored before estimates of operational usefulness can be made of any of the study findings. Consequently, scoring methods developed to assess interpreter performance using projected displays would have to be applicable to performance with more conventional (light table) displays.

Finally, a fundamental problem exists in determining what is a "correct" response, a consideration basic to the establishment of accuracy or completeness or efficiency scores. Mention has already been made of the determination of "ground truth" which is the first necessary step in criterion development. Attention is now directed to the concept of the hierarchical target family structure which provides a scoring rationale.

For purposes of the present study, a rather elemental weighting system was employed. The system assigned unit values to major target family responses, two point values for responses at the second descriptive level and three point values for responses at the third descriptive level. This scoring scheme achieved the purpose of giving credit for detection of the presumably more subtle details that permitted greater precision in identification. However, the weighting system must be considered as unrefined because of the unsupported assumption of equal difficulty of identification among targets at the same level in the hierarchical structure.

In determining specific target point values (see Table 1-3) the following guide may be used. All targets prefaced with Roman numerals are of one point value; targets having Arabic numerals are of two point value; and targets with lower-case alphabetic designators are of three point value. Thus, all "right" and "wrong" scores used in this study are weighted scores, and are referred to hereinafter as "points."

The criterion measures which were employed during the study are cited below. In certain instances, where the definition may not be obvious, descriptions of computational methods have been added.

- Time
- Number of "Right" (weighted) response points
- Number of "Wrong" (weighted) response points: Points assigned for reports of targets which were not present in the image. This was a measure of errors of commission. Errors of omission were obtained by subtracting the "Right" response total from total points possible as determined by "ground truth" scoring efforts.
- Accuracy: Accuracy was defined as that percentage of the subject's total responses which were correct; found by computing the ratio of correct response points to total points produced.

$$\frac{\text{Right}}{\text{Right + Wrong}}$$

- Completeness: Completeness was defined as the ratio of correct response points to the total point value of the imagery. Erroneous responses, which were considered in the accuracy score, are omitted from this formula.

$$\frac{\text{Total Right Points}}{\text{Total Possible Pts.}}$$

- Efficiency: Since the project was directed toward training rapid perceptual skills, a measure of the subject's speed in identifying targets was needed. Time alone, however, is not meaningful without reference to some other aspect of performance. Consequently, an "efficiency index" was defined which expressed the number of correct response points per unit time. It was defined as the ratio of correct response points to the total time spent viewing the display.

$$\frac{\text{Total Right Points}}{\text{Total Time (sec.)}}$$

- Effectiveness: A score occasionally employed in graphic comparisons of group performance. Obtained by dividing the mean Completeness score by the mean time spent viewing the images.

$$\frac{\text{Completeness}}{\text{Time}}$$

A final criterion measure which was employed was concerned with the measurement of the amount of information transmitted, and the information processing rate of the subjects. The application of information theory techniques to the present study appeared to afford insights into study results which might be less strikingly evidenced by conventional analytic methods.

The several criterion measures described above were applied to subjects' responses to the 90 images of the Projected Display test and to the 17 three-frame sequences of the Light Table test. These two tests, and their related scoring methods, formed the criteria against which the adequacy of the training program was assessed.

C. THE EQUIPMENT

The projection and display equipment consisted of the following items:

1. Automatic 35mm Slide Projector (Wollensak Model P-815)
2. Harvard Electric Shutter (3-inch aperture)
3. Electronic Photo-Timer (Lektra Labs, Inc. Model TM5)
4. Pola-Coat Rear Projection Screen (20"x20")
5. Slide Advance/Repeat Remote Control Switch Box
6. Motorized Light Table

The slide projector was modified to incorporate the Harvard electric shutter and the timer, thus providing tachistoscopic capability. The timer permitted the exposure duration to take values from 0.1-second to 40.0-seconds at 0.1-second intervals while a timing device incorporated in the projector provided for a choice of inter-slide intervals.

Subject testing in the final stages was conducted on an individual basis. Figure 1-1 shows the equipment configuration for the testing situation. The slide advance/repeat control was a remote control device which the subject used either to advance to a new slide or to request an unlimited number of repeat projections of the slide then positioned in the projector.

The Pola-Coat rear projection screen measured 20"x20" and was mounted vertically, approximately 30 inches from the subject.

The image display size on the screen was selected on the basis of a preliminary test which explored subject response accuracy across the target family spectrum for 10-inch and 20-inch display sizes. It was found that nearly twice as many targets were correctly identified at the 20-inch display size than at the 10-inch display size. Apparently, the advantage

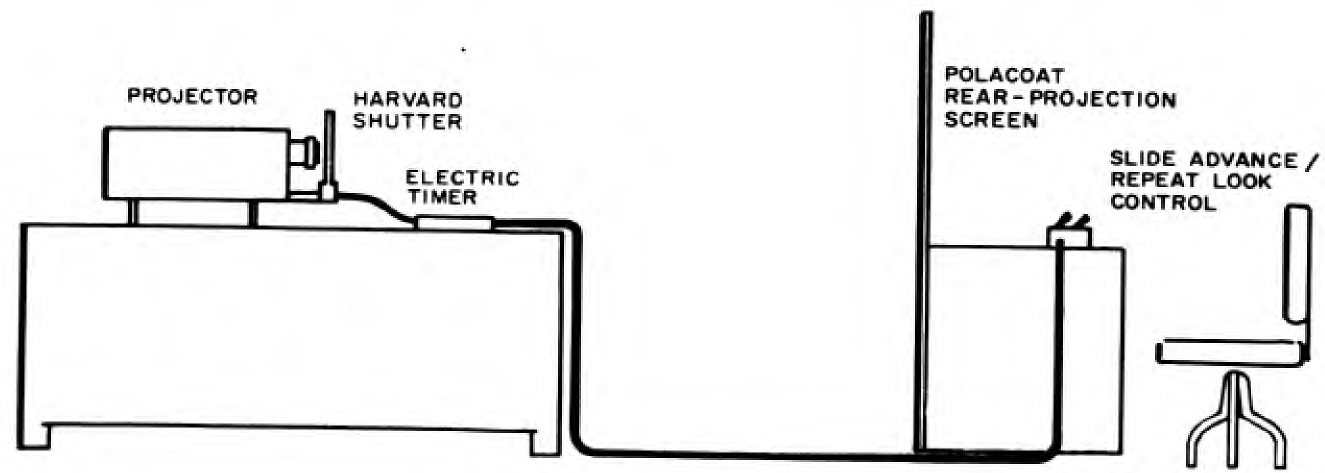


FIGURE 1-1. Equipment Configuration during the Testing Stage.

gained by sharper images at the smaller display size was more than offset by the advantage of the larger magnification.

A motorized light table was the final item of equipment employed. Electrically driven, the equipment allowed for the display of aerial roll film up to a 9x9-inch size. The central display area of the table was a three-panel section of 3/4-inch ground glass, each panel 9x9-inches, over which the film was transported. Illumination was provided by six 15-watt fluorescent bulbs mounted below the ground glass (two per panel). The speed of film transport and direction of movement was controlled by a lever which provided for three speeds in either the film advance or rewind direction.

D. THE EXPERIMENT

1. Subjects. The 30 subjects used in evaluating the training program were drawn primarily from local installations. Nearly all were currently engaged in interpretation work, most had received formal interpretation training in military schools, and the median interpretation experience level exceeded four years.

The subjects were administered a brief screening questionnaire of untested reliability in an effort to derive some objective basis for equating experimental and control groups. The "test" consisted of 15 photographic prints selected to represent a variety of man-made features. The prints were assembled on a single sheet and reproduced using a Xerox copy machine. The process substantially degraded the print quality, resulting in target images which were extremely difficult for unskilled subjects to identify. The test was administered to a variety of persons in addition to the test subjects and the range of scores achieved was fairly substantial. Photointerpreters (including the test subjects) scored among the highest;

persons not formally trained in photointerpretation but with experience in intelligence work scored next highest; and persons with neither interpretation training nor intelligence experience scored lowest. The testing approach, though far from producing a refined and precise testing instrument, appeared to have some ("face") validity and, therefore, some merit as a means of objectively equating the experimental and control groups.

The subjects for this study were obtained from a variety of sources. Of the thirty subjects, 16 were from the Naval Photographic Interpretation Center in Suitland, Maryland; four were from the RCA Service Company in Alexandria, Virginia; four were RCA, Data Systems Center personnel, and six were from various private and governmental agencies in the Washington area. The thirty experienced photointerpreters were divided into two groups of fifteen subjects each, largely on the basis of scores attained on the screening test, and these served as the Experimental and Control groups. Table 1-4 presents descriptive information for both groups.

2. The Training Curriculum. Summary flow diagrams of the major steps accomplished during the training program for the Experimental and Control groups appear as Figures 1-2a and 1-2b.

The training program covered a one-week period and was administered in its entirety to the 15 subjects who served as the Experimental group. The Control group received an abbreviated version which provided sufficient information and practice to allow them to employ approved target terminology and to understand testing procedure. Control group "training" and testing were accomplished during a two-day period.

TABLE 1-4

DESCRIPTIVE CHARACTERISTICS OF THE EXPERIMENTAL AND CONTROL GROUP SUBJECTS

	Experimental (N-15)	Control (N-15)
Median Age	38.0	34.0
Type Formal PI Training		
Navy	4	11
Army	3	2
Air Force	1	-
Marine	1	-
Other (Civ.)	4	-
None (OJT)	2	2
Median Experience Level (Years Estimate)*	4.5	4.9
Median Current PI Work Load (Hrs./Week Estimate)*	5.0	4.0
Median PI Questionnaire Performance		
(1) Accuracy*	51.0%	49.5%
(2) Completeness*	56.0%	54.0%

* Chi-square test conducted, dividing distributions at median values. Differences between groups not statistically significant.

Space limitations precluded training either group as a whole; each group was divided into two approximately equal subgroups and training was administered accordingly.

A detailed outline of the Experimental group daily training activities is presented below.

RAPID IDENTIFICATION TRAINING PROGRAM
(Experimental Group)

MONDAY

I. Study Explanation: an introduction to the main problem and purpose of the study; the approach taken to reach the training goal; a demonstration of the equipment;

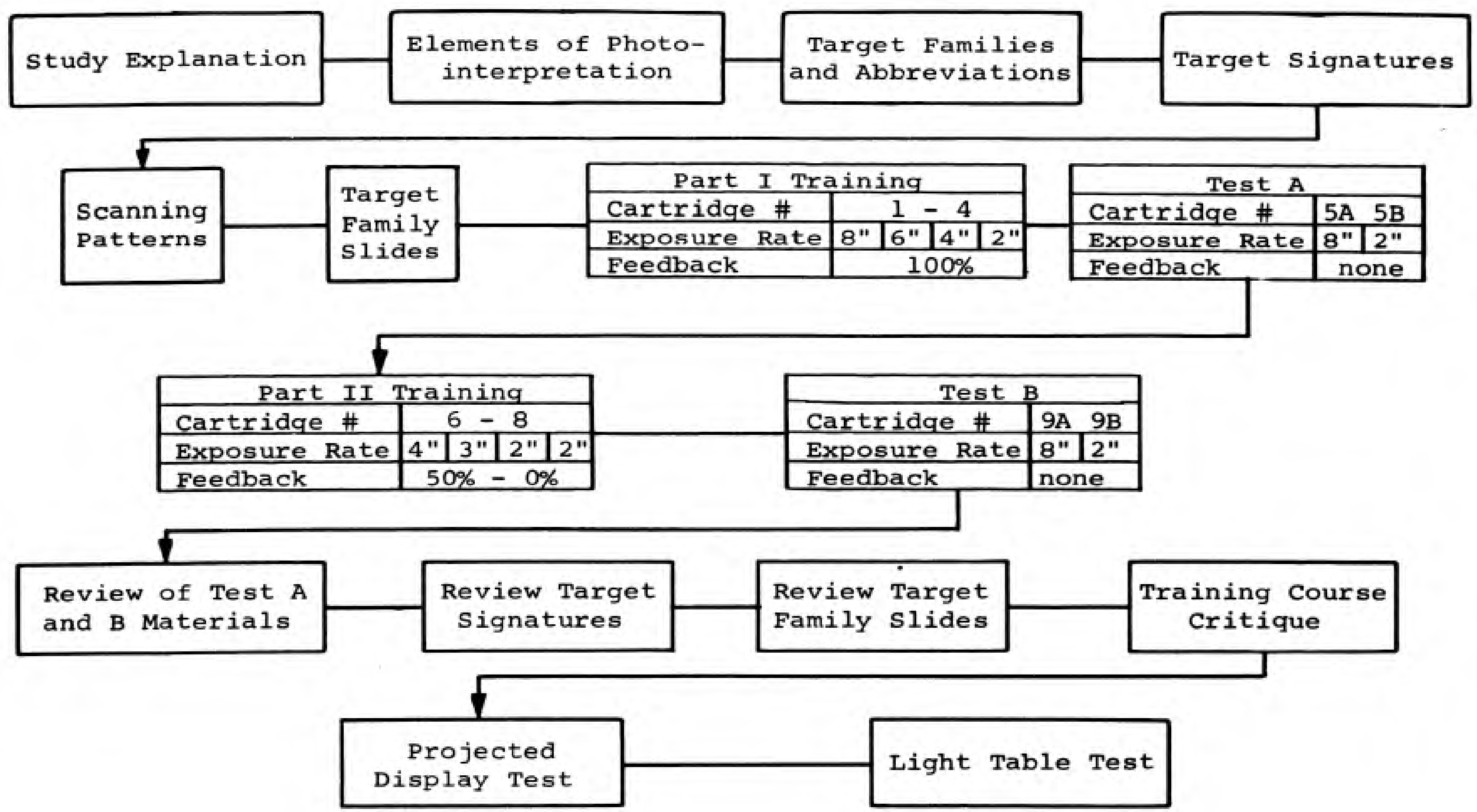


FIGURE 1-2a. Flow Diagram of Training Course Elements for the Experimental Group.

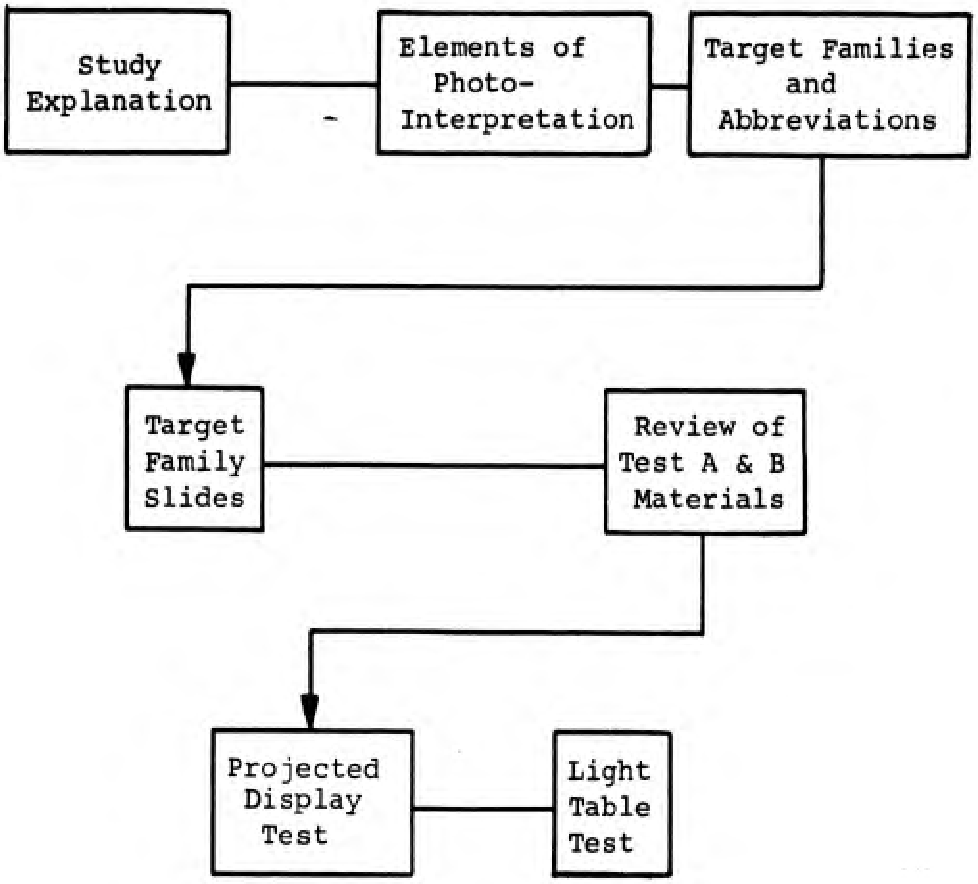


FIGURE 1-2b. Flow Diagram of Training Course Elements for the Control Group.

and a preview of the training to be covered during the week.

II. Elements of Photointerpretation: a brief review of the detection, identification, and analysis stages comprising the photointerpretation process with emphasis upon the first two stages.

III. Target Families and Abbreviations: the 42 target names are introduced along with the standardized abbreviations; practice developing target name/abbreviation associations and familiarity with the target family structure; explain and practice using the hierarchical reporting system.

IV. Target Signatures: the concept of target signatures is defined and discussed; useful sources of signature elements are covered along with an introduction to speed signatures.

V. Target Signature Illustrations and Descriptive Information: each target is discussed in terms of its complete signature, speed signature, and associational elements; signature illustrations are provided for all targets.

TUESDAY

VI. Scanning Patterns: typical scanning errors are mentioned; procedures for more effective scanning behavior are presented and practiced.

VII. Target Family Slides: actual slides representing all of the targets are shown along with an explanation of the presentation sequence.

VIII. Part I Training: 4 cartridges containing real slides and feedback slides are used; feedback consists of following a real slide with a synthetic one which contains signature drawings and abbreviations of the target(s) present in the real slide; after the synthetic slide is exposed, the actual image is once again displayed briefly; 100% feedback used; cartridges are presented four times each at 8-seconds, 6-, 4-, and 2-seconds; subjects record their answers.

IX. Test A: used to determine improvement in rapid identification ability. Subjects tested with a new slide series at 8-seconds and a series at 2-seconds; no feedback.

WEDNESDAY

X. Part II Training: procedure here is essentially same as Part I Training except: 1) feedback begins at 50%, reduces to 25%, and then decreases to 0%, 2) exposure durations begin at 4-seconds and are reduced to 2-seconds. New slides are used in Part II Training.

XI. Test B: used to determine improvement in rapid identification ability. Subjects tested with a new slide series at 8-seconds and a series at 2-seconds; no feedback.

THURSDAY

XII. Review of Tests A and B: test slides are reviewed at 2-second exposure to provide verbal feedback on previously unreinforced imagery.

XIII. Review/Critique of Target Signatures: target signature training points are reviewed prior to final testing; subjects' comments on adequacy of signatures in representing targets are obtained.

XIV. Review Target Family Slides: a review of the specific targets used in the study is provided at a 2-second exposure duration for each image.

XV. Written Critique: the subjects' comments on selected portions of the training are solicited.

FRIDAY

XVI. Final Testing: Projected Display test involving 3 cartridges of 30 slides each; subject controls number of exposures required for each slide; Light Table test involving motorized light table and 9x9-inch aerial roll film positive transparency.

The treatment accorded the Control group required a compromise between the goal of providing at least equal exposure to the quantity of training slides viewed by the Experimental group, and avoidance of any rapid identification training.

Since the Control group could not view all the slides without being exposed to the same brief slide presentation durations, it was necessary to use a smaller sample of slides and present them for longer periods. The task thus became one of selecting slides representative of those included in the final test and exposing these for a time equivalent to the total slide viewing time used by the Experimental group. Because of the level of sophistication of the subjects, it was felt that this scheme would not be detrimental. All subjects were (because of their previous experience) familiar with the targets of interest, and the main purposes of this portion of the training were to acquaint them with the limited list of targets (42) and the terms they were to use in reporting detected targets.

The Control group did not, of course, receive target signature training, scanning training, or controlled feedback on their responses to displayed imagery. Hence, their total training time was substantially shorter than that of the Experimental group.

They were accorded a full explanation of the study, and were drilled on the list of targets and acceptable abbreviations of the target names. They were exposed to a majority of the same slides used in training the Experimental group (specifically, the target family slides and the material comprising Test A and B); the main difference being that at least 30-seconds was allowed for each exposure. Thus, the Control and Experimental groups viewed highly similar training imagery for approximately the same total amount of time.

An outline of the Control group training program is presented below. Where a training item was exactly the same for both groups, it is so noted.

RAPID IDENTIFICATION TRAINING PROGRAM/Control Group

MONDAY

I. Study Explanation: SAME

II. Elements of Photointerpretation: SAME

III. Target Families and Abbreviations: SAME

IV. Target Family Slides: SAME

V. Review of Tests A and B Material: provides control subjects with equal exposure to similar imagery seen by Experimental group without benefit of rapid viewing practice and feedback slides.

TUESDAY

VI. Final Testing: SAME

E. CONDUCT OF THE TEST

For the Projected Display Test, the subject was seated approximately 30 inches from the rear projection screen. The screen was positioned so that the center was approximately at eye level. The subject was instructed in the use of the remote control, slide advance/repeat box and was permitted to practice. The 30 slide cartridge of Large Scale imagery was loaded in the projector and the timer was set for a 2-second exposure duration. The subject pressed the slide advance toggle and reported his identifications verbally. If the subject desired a repeat exposure of the same slide he pressed a pushbutton and the slide was redisplayed for 2-seconds. The experimenter recorded the verbal responses and the number of exposures per slide.

For the Light Table Test, the subject was provided only with a record form for listing his identifications - no mechanical aids were permitted. Starting and stopping times were recorded.

F. SCORING

Concurrently with the conduct of the training program, work began on the preparation of a computer program to permit electronic data processing of the subjects' responses. The scoring rationale was developed, based upon the hierarchical scheme previously noted, and instructions were prepared delineating the various summary measures which were to be derived.

The subjects' responses were entered on cards and converted to tape records. The RCA 301 computer performed the complete scoring of responses during training (Experimental group Part I and Part II training, Tests A and B) as well as both groups' responses to the Projected Display test imagery. A computer program was prepared especially for this scoring operation. Later statistical analyses were, however, accomplished with standard statistical programs available for all RCA electronic data processing equipments.

The Light Table Test was hand scored, and manual tabulations were made of responses for the analyses concerned with Accuracy, Completeness, and Efficiency. Apart from these, the bulk of the scoring and all of the correlational analyses were machine-accomplished.

G. RESULTS

The findings of a training experiment such as the present study may be treated in a variety of ways. First, of course, is the issue of post-training differences between the Experimental and Control groups. These differences may be assessed in terms of several criterion measures, as was previously discussed. Among such measures are indexes of accuracy, completeness "efficiency," and time; and measures of information processing and information processing rate; all of which may be applied to subject responses under two different image display conditions.

The second area of interest concerns the nature of the criterion measures themselves; their respective uniqueness and their reliabilities.

An additional area of interest, provided that significant group differences obtain and that these are discovered for unrelated criterion measures, concerns the possibility of identifying which aspects of the training program were largely responsible for the group differences. A related question concerns the possibility of identifying where in the training cycle the discovered improvements occurred.

Finally, and again assuming the existence of significant group differences, there is a question of whether available biographical or pretest measures might serve as predictors of the criterion measures.

These and other issues are treated in the various analyses which are presented below.

1. Group Comparisons: Projected Display. Comparisons were made of Experimental and Control group performance, in terms of the several criterion measures, with respect to achievement on large, medium, and small scale imagery. In many instances the target signatures logically appear to be at least partly a function of scale. It was felt, therefore, that training benefits might accrue differentially for imagery of differing scale values.

Since the intent of the training was to promote rapid viewing, one of the first analyses was concerned with the mean time spent in viewing each slide. The results which are graphically portrayed in Figure 1-3 indicate that, on the average, the Experimental group members spent about one-half as much

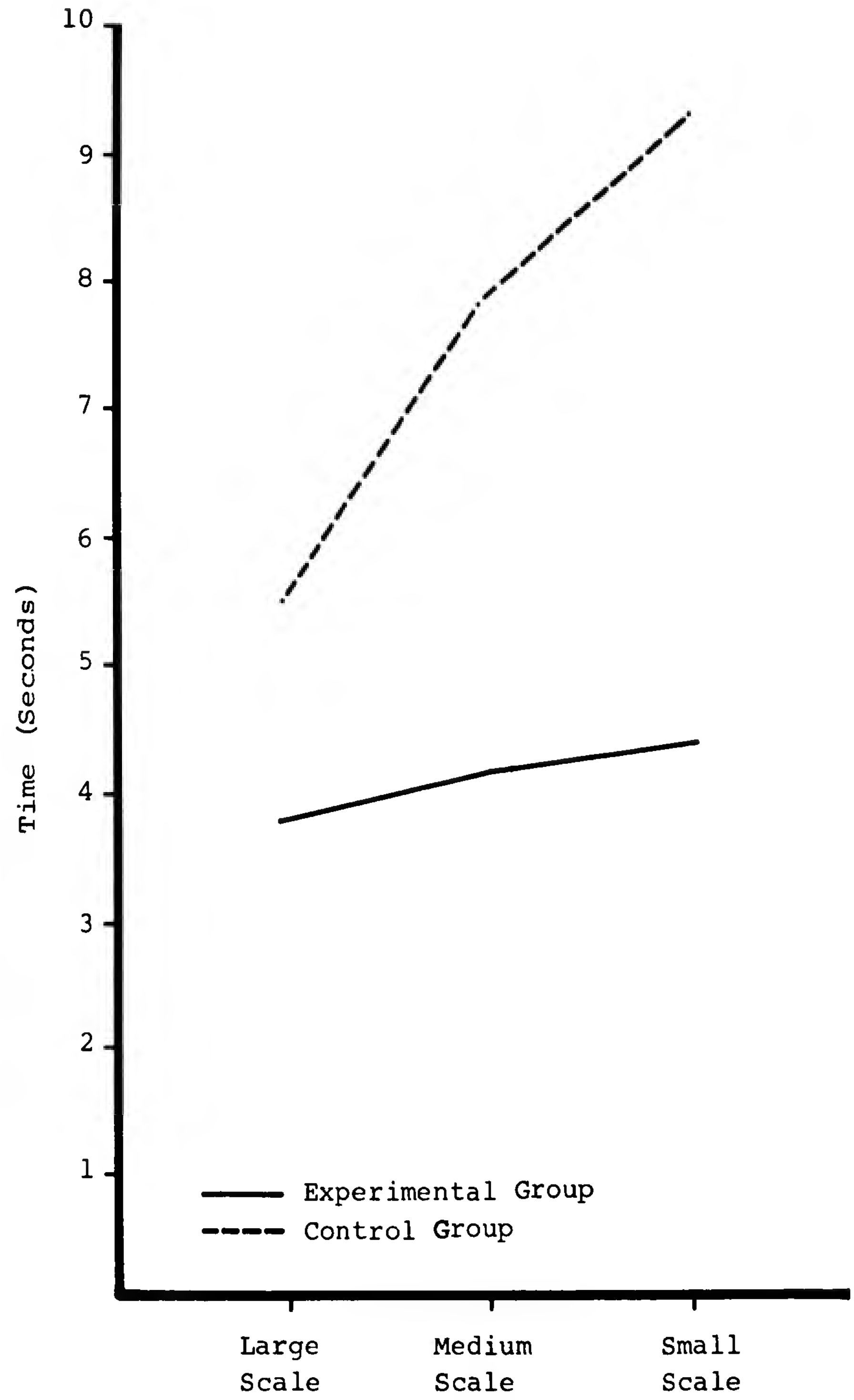


FIGURE 1-3. Comparison of Experimental and Control Groups in Terms of Mean Times Spent Viewing Imagery at Different Scale Values - Projected Display Test.

time per slide as did the Control group members. The effect of image scale is also quite pronounced as evidenced by the nearly four seconds per slide greater time spent by the Control group. The Experimental group, on the other hand, spent an average of only one-half second per slide longer while viewing small scale imagery as they did while viewing large scale displays. The group time differences at each scale value were statistically significant beyond the .01 level of confidence.* Provided that the measures of accuracy and completeness for the two groups are similar, this would suggest an impressive gain in interpreter performance which could be directly attributable to the training program. Consequently, attention was directed toward the calculation of other criterion values.

There are four conditions possible in scoring an image interpreter's responses. A target may be present and be reported as present; a target may be present but not reported as being present (error of omission); a target may not be present but is reported as being present (error of commission); or a target may not be present and is not reported as being present. Although "zero-target" imagery was employed during training, none was used during testing. Consequently, the fourth condition may be temporarily regarded as irrelevant. Tabulations of responses subsumed under the other three conditions were made for both Experimental and Control groups, and these appear in Table 1-5. It will be noted that at all image scale values the Experimental group achieved a greater percentage of correct responses, had fewer error of omission, and (at medium and small scale values) had more errors of commission.

---

*Unless otherwise noted, significance tests are Chi-square values with 1 degree of freedom, based upon a division of the respective distribution at the median value.

For subsequent analytic purposes, responses were multiplied by target weighting values. Those responses contained in columns A and C of Table 1-5 were, when properly weighted, used to calculate Accuracy scores, and responses in columns A and B, when properly weighted, were used to calculate Completeness scores.

TABLE 1-5

COMPARISON OF EXPERIMENTAL AND CONTROL GROUPS IN TERMS OF CORRECTNESS, ERRORS OF OMISSION, AND ERRORS OF COMMISSION

(Projected Display Test)

	Column A Target Present Reported Present	Column B Target Present Not Reported	Column C Target Absent Reported Present	(N) Total Responses	(T) Total Time (Sec)
Large Scale					
Experimental	64.7%	24.3%	10.9%	893	1716
Control	57.5%	32.3%	10.0%	884	2494
Medium Scale					
Experimental	58.7%	29.0%	12.2%	1333	1898
Control	54.1%	37.4%	8.4%	1278	3574
Small Scale					
Experimental	58.2%	31.7%	10.0%	1400	1920
Control	54.9%	36.7%	8.2%	1373	4180

Comparison of group performance in terms of the measures of Accuracy and Completeness revealed additional training benefits beyond that of time improvement. The Experimental group achieved significantly higher Accuracy scores on large scale imagery and significantly higher Completeness scores on both large and medium scale imagery. The levels of significance achieved were .05, .01, and .01, respectively. None of the remaining differences was statistically significant. The Accuracy and

Completeness scores are graphically presented in Figure 1-4.

When the groups were compared in terms of "efficiency" (defined as the number of correct response points per second of image display) the differences were striking. The Experimental group scored significantly (.01 level) higher than did the Control group at each of the image scale groupings (see Figure 1-5).

Finally, a check was made of the Completeness scores achieved by the Experimental and Control groups in terms of the major target family categories. The intent of this analysis was to determine whether the Experimental group's improvement was general across target families (in which case the training benefits could be largely attributable to the speed emphasis and scanning training) or whether improvement was specific to target families (in which case at least some credit for the improvement would be attributable to the target signature training).

Median Completeness scores were calculated for the two groups in terms of each major target family category. The Experimental group median Completeness score was divided by that of the Control group in order to express these as a ratio (E/C Ratio) of relative performance:

Target Family	E/C Ratio
Military	120
Industry	149
Rail	111
Airfield	100
Port/Harbor	100
Power	140

The results suggested (a) that the signature training may be at least partially responsible for the Experimental group's superior performance, and (b) that not all signatures are functioning

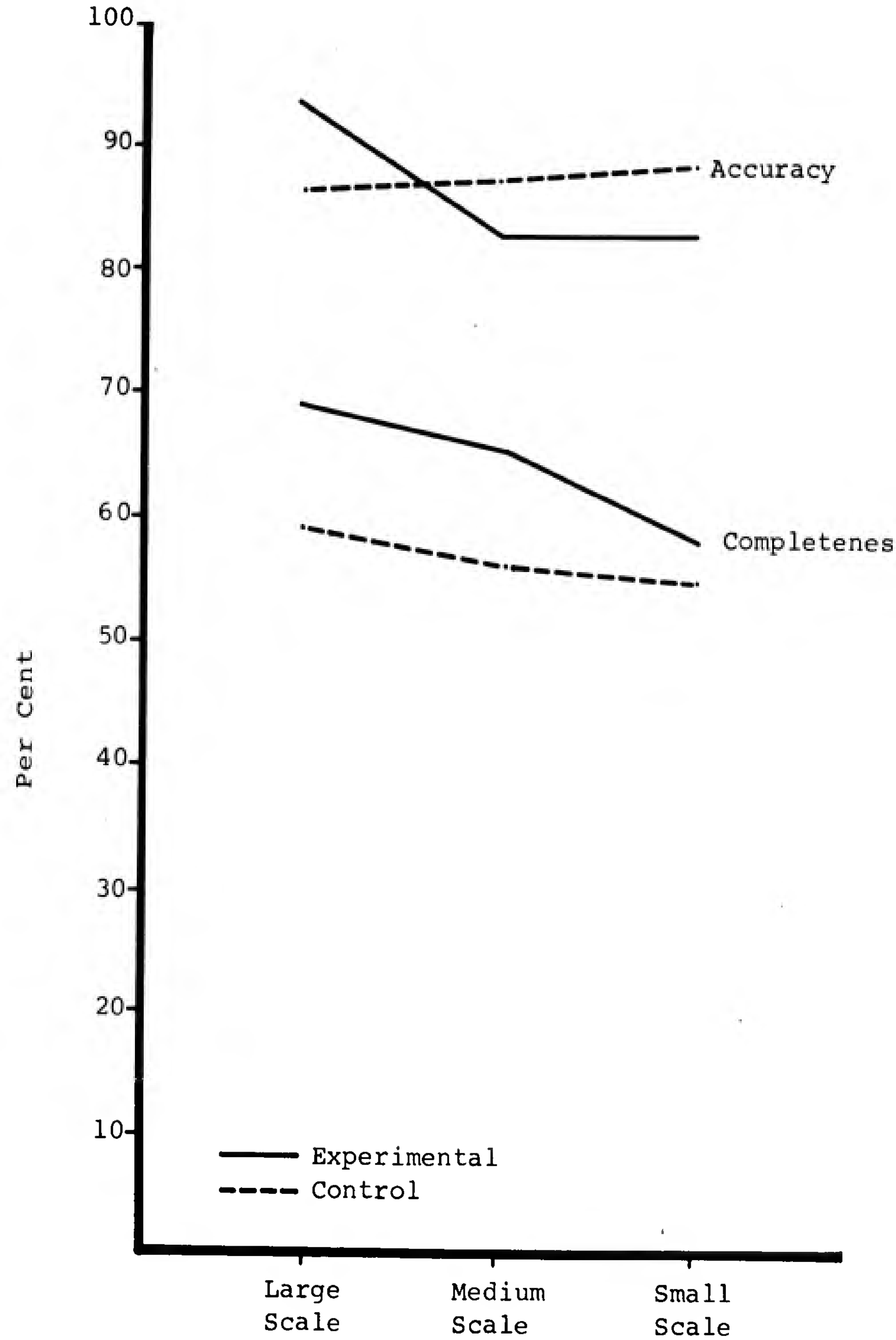


FIGURE 1-4. Comparison of Experimental and Control Groups in Terms of Accuracy and Completeness Scores by Image Scale - Projected Display Test.

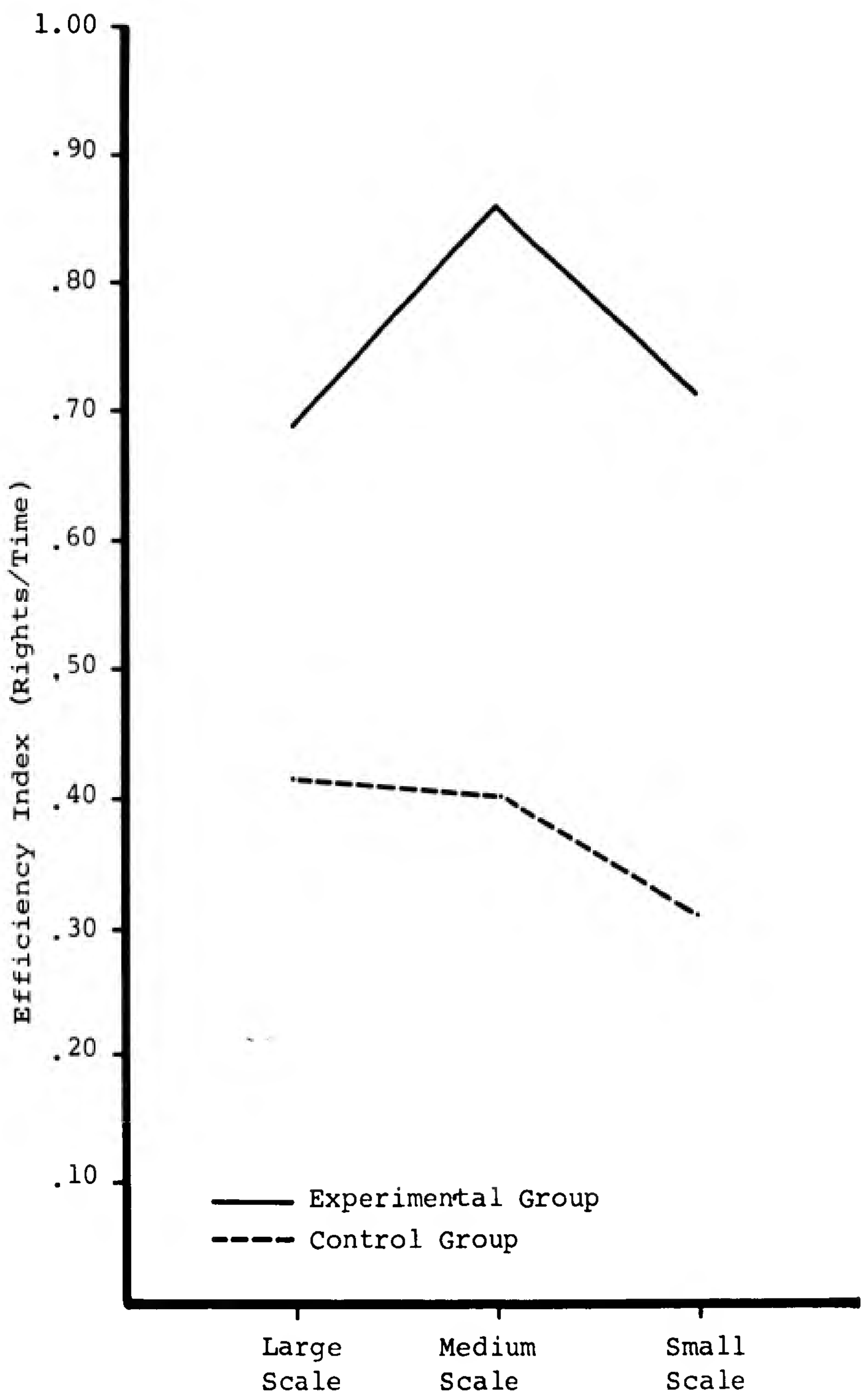


FIGURE 1-5. Comparison of Experimental and Control Groups in Terms of Efficiency Scores by Image Scale - Projected Display Test.

equally well in promoting this improved performance level. No differences are to be noted in Completeness scores for Airfield or Port/Harbor categories, and only modest gains appear in the Rail category. As a consequence of these sorts of analyses, the Appendix to this report contains signature data limited to those target types for which Experimental group performance was measurably superior. It would be misleading, and possibly harmful, to include non-functioning graphic or verbal signature materials.

2. Group Comparisons: Light Table Display. As was previously mentioned, the rear projection display represented a departure from the customary viewing mode for photointerpreters. There was, consequently, a concern that the Experimental group might have established a "time set" unique to the viewing condition and that they might revert to slower viewing habits when tested with the more conventional (light table) display. It was found that the Experimental group did, in fact, revert to a slower work pace, but was still substantially more rapid than was the Control group.

With respect to the criterion of time, the Experimental group performed the task significantly more rapidly than did the Control group. (Mean times of 27.6 minutes and 36.5 minutes respectively; difference significant beyond the .001 level of confidence.) There were no significant differences between the groups for measures of Accuracy or Completeness at any of the image scale groupings. However, because of the substantial time difference in task accomplishment, the Experimental group scored significantly higher than did the Control group in terms of the Efficiency Index (significant beyond the .02 level of confidence).

3. Human Information Handling. The 43 target family categories (42 target names plus a "no target" category) considered as a nominal scale, may be organized in matrix format for information analysis purposes. In the present instance, the keyed targets for each slide were entered as the transmitted stimuli and, for each stimulus, the subjects' responses were recorded.

The data were analyzed separately for Experimental and Control group subjects for each of the three categories of image scale. The analysis was accomplished only for the Projected Display Test.

The information presented to the subject (per stimulus exposure) was calculated from the marginal distribution since categories of stimuli (target categories) were not presented equally often.

The computational method was straightforward (Attneave, 1959) and yielded data concerning group information handling performance which were highly similar to those resulting from the above-noted analyses. For example, the amount of information, in bits, transmitted from stimulus to response was as follows:

	Large Scale	Medium Scale	Small Scale
Experimental Group	2.5	2.2	2.1
Control Group	2.2	1.9	2.1

Comparing these with an estimated information per stimulus of 4.0 bits yields curves closely approximating the Completeness curves in Figure 1-4.

An approximation of information processing rate may be obtained by dividing the mean information transmission value per stimulus-response unit by the mean number of seconds the displays were viewed. This is not a true measure of human information processing rate since the response times are not included.

However, assuming an equivalent average response time for the two groups, the substantial difference in relative performance of the two groups can be appreciated from the following values (expressed in bits per second):

	Large Scale	Medium Scale	Small Scale
Experimental Group	1.29	1.50	1.53
Control Group	.78	.69	.69

4. Intercorrelations Among Criterion Measures. In an earlier section, mention was made of the logical independence of such aspects of interpreter performance as measures of Accuracy, Completeness, and Efficiency. The hypothesis that these were indeed independent was assessed during the data analysis portion of the study. In addition, correlational analyses were conducted relating performance on the Projected Display Test and the Light Table Test. The concern was that these tests were in fact measuring similar aspects of interpreter performance. Finally, correlational analyses were conducted of the relationships of the image scale subtest scores for each of the criterion measures. These relationships may be interpreted in terms of internal consistency estimates of test reliability.

a. Independence of Criterion Measures. On both the Projected Display Test and the Light Table Test the measures of Completeness-Efficiency, and Accuracy-Efficiency were uncorrelated.* On both tests the measures of Accuracy-Completeness were negatively related (-.56 and -.43, respectively; correlations significant beyond the .01 and .05 levels of confidence, respectively).

*Unless otherwise noted, correlational analyses cited refer to product moment coefficients.

b. Projected Display vs. Light Table Test. Three of the four criterion measures were significantly (.01 level) related to their counterpart measures on these two tests. The several coefficients were as follows:

Accuracy	.54
Completeness	.77
Efficiency	.65
Time	.35

c. Criterion Measure Stability Across Image Scale Subtests. Thirty-two of the 42 intercorrelations were sufficiently high to warrant the conclusion that the majority of subtest measures were significantly related to their counterpart measures. Exceptions were the Accuracy scores for both the Projected Display and the Light Table large scale subtest. These tests did not relate significantly to other subtest Accuracy scores.

The intercorrelations for the various criterion measures are presented in Table 1-6.

5. Training Course Effectiveness. During conduct of the training course a number of brief tests were administered to the Experimental group. It was intended that these would serve as a measure of possible progress so that, if differences were found to exist between Experimental and Control groups, it might be possible to identify the amount of training required to produce such differences.

Figures 1-6a and 1-6b summarize the results of several of the testing efforts which are directly related.* Test 1 refers to

*Since cartridges were repeatedly displayed and since exposure durations were reduced during successive presentations, certain sections of the records are either contaminated because of prior exposure, or not comparable because of exposure time differences. These sections were not included in the analyses.

TABLE 1-6

INTERCORRELATIONS AMONG PROJECTED DISPLAY (PD) AND
LIGHT TABLE (LT) SUBTESTS FOR EACH CRITERION MEASURE*

	Projected Display		Light Table		
	Med.	Sm.	Lg.	Med.	Sm.
Accuracy					
PD Test					
Large	.41	.23	-.06	-.03	.16
Medium		.63	.35	.55	.48
Small			.21	.62	.41
LT Test					
Large				.43	.38
Medium					.62

	Projected Display		Light Table		
	Med.	Sm.	Lg.	Med.	Sm.
Completeness					
PD Test					
Large	.71	.60	.63	.50	.78
Medium		.77	.53	.51	.68
Small			.28	.63	.65
LT Test					
Large				.34	.54
Medium					.65

	Projected Display		Light Table
	Med.	Sm.	Total
Efficiency			
PD Test			
Large	.82	.69	.53
Medium		.89	.59
Small			.67

	Projected Display		Light Table
	Med.	Sm.	Total
Time**			
PD Test			
Large	.74	.66	.39
Medium		.97	.33
Small			.31

* $r = .36, p < .05$ $r = .47, p < .01$

** Time values for completing portions of the Light Table Test were not obtained. Only the total time was recorded.

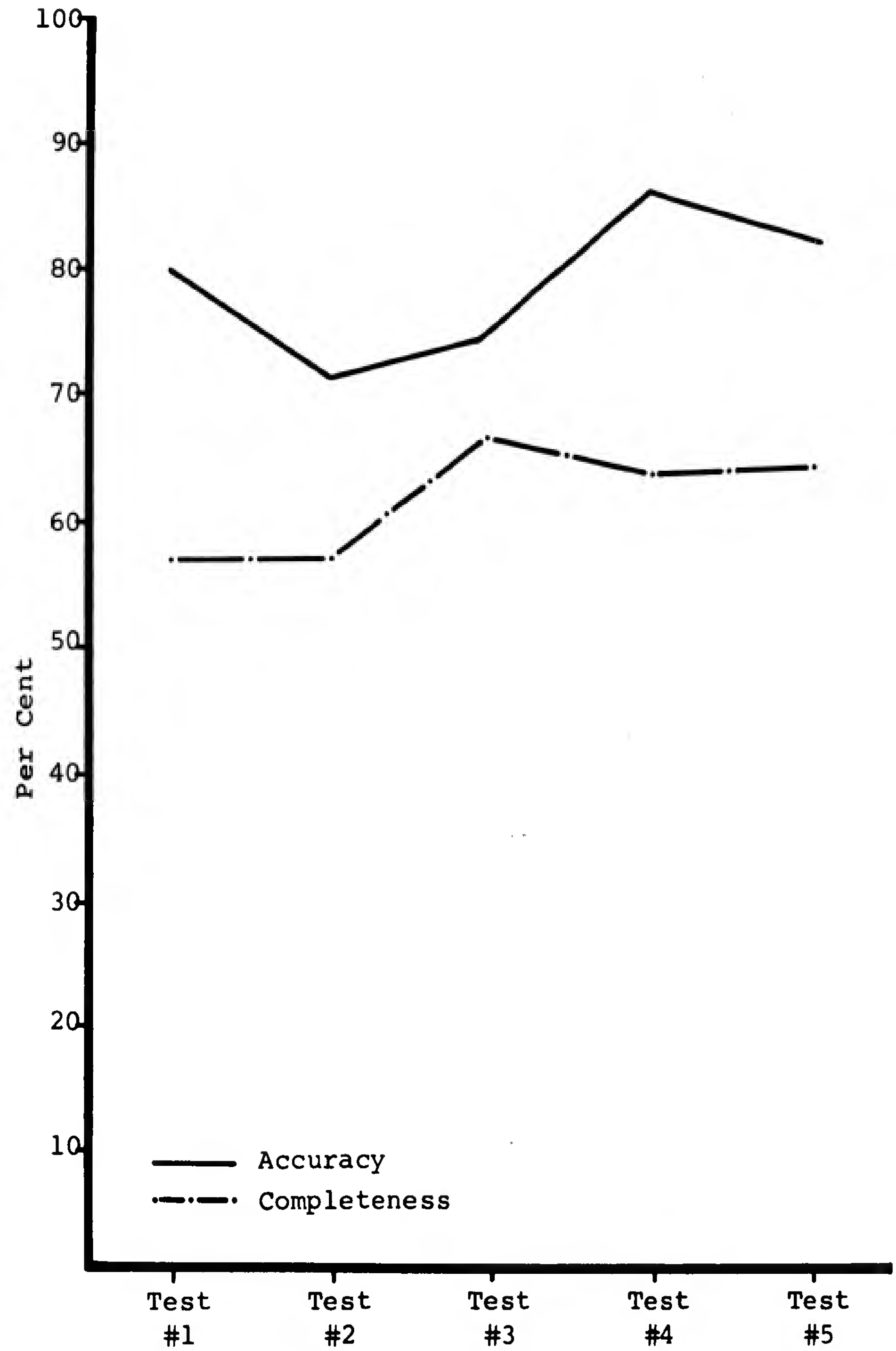


FIGURE 1-6a. Accuracy and Completeness Scores Achieved during Conduct of Experimental Group Training/Testing.

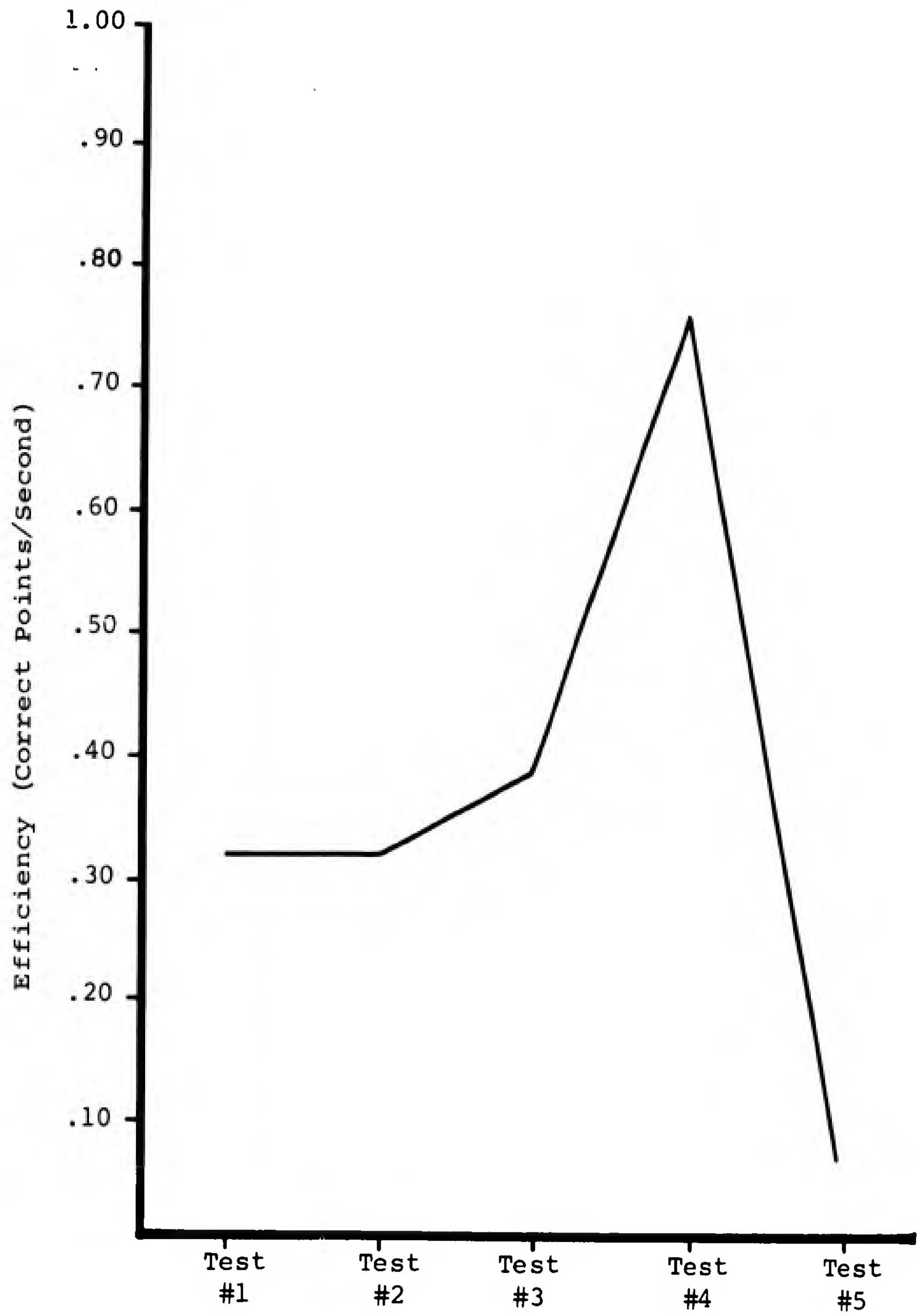


FIGURE 1-6b. Efficiency Scores Achieved During Conduct of Experimental Group Training/Testing.

performance on the first (eight-second exposure) administration of the four slide cartridges comprising Part I training materials. The entries for Test 2 and 3 refer only to the slides exposed for eight seconds; hence, scores on these are directly comparable to the Test 1 measure. The Projected Display (Test 4) and the Light Table (Test 5) scores are, of course, based upon a subject-determined exposure time.

The Projected Display (Test 4) scores are all significantly (.01 level) greater than the Test 1 scores and, for the Accuracy and Efficiency measures, significantly (.01 level) higher than the Test 3 scores. For the Completeness measure, the Projected Display Test score is significantly (.01 level) higher than the Test 2 score but is not significantly higher than the Test 3 score. The extremely rapid rise in the Efficiency score from Test 3 to the Projected Display Test suggests that the 8-second machine-paced display time was inordinately long, at least after the initial block of training imagery (imagery presented prior to Test 3 administration).

The Light Table (Test 5) scores of Accuracy and Completeness are similar to those obtained on the Projected Display test. However, the Light Table test Efficiency score is significantly different from all other Efficiency scores (.001 level of confidence). This substantial lowering in Efficiency would suggest that a continuous display may encourage a subject to view an image for a much longer time than is necessary. The longer viewing time did not increase either Accuracy or Completeness scores.

6. Biographical and Pre-test Data. As noted in Table 1-4 certain biographical and pre-experiment test data were obtained for subjects participating in the research study. During the data analysis phase, each of the sources of potential predictor

information was assessed in terms of its relationship to later measures of Accuracy, Completeness, Efficiency and Time on each of the two display format tests. None of the variables of experience, prior training, or pre-experiment test scores was found to be significantly related to any of the criterion measures.

SECTION 2

IMAGE AND DISPLAY CHARACTERISTICS INFLUENCING INTERPRETER PERFORMANCE

A. INTRODUCTION

Although the summary analyses of interpreter performance in terms of Accuracy, Completeness and Time provide useful information, this knowledge concerns but one component in the system which also includes the nature of the stimulus material. Image characteristics and modes of displaying the images place limits upon the amount of information that can be extracted from a pictorial image, and these characteristics interact with the interpreter's sensory and cognitive abilities in irregular fashion.

The fact that the specially trained group performed more rapidly and more accurately than did the group which did not receive specialized instruction is of obvious importance. The findings are relevant to a host of interpretation task assignments, from routine screening of photography to real time interpretation in aerial reconnaissance vehicles.

Of at least equal importance, however, is the analysis of subject performance in terms of characteristics of the imagery itself. Was the time gain achieved solely because of an increased work tempo - a machine-paced "pressure" specific to the test situation but one which would not be expected to hold up in an operational setting? Were gains achieved across the spectrum of target types, which might suggest the work tempo hypothesis is correct? Or were gains specific to target types, which would tend to support the hypothesis that the "target signature" abstractions had indeed contributed materially (but differentially) to interpretation speed and accuracy?

What is the relation between target size or rated conspicuity to the probability of detection, and did the trained subjects tend to excel in picking out small target patterns as a function of their prior training in graphic signatures? What are the implications of the study findings to the issues of interpreter training and efficient operational usage?

Conduct of the empirical assessment of the training program provided an opportunity to study the possible influences on interpreter performance of many variables which bear upon the above questions.

There are several reasons, both theoretical and practical, for concern with testing such relationships, and these are listed below, not necessarily in their order of importance. First, there is always a concern for learning more about how persons perceive and identify objects of interest. The literature on visual perceptual research is enormous but the laboratory findings are often not easily applied to operational problems. Knowledge of the threshold angle for perceiving a Landoldt circle under specified conditions of illumination does not aid in selecting the image scale needed to achieve a 90 per cent probability of detecting a medium tank in a wheat field.

Second, although many of the investigated relationships exceeded the initially proposed study scope, their inclusion served as an aid in explaining the significance of the findings of the central study.

Third, the analyses were conducted in some cases simply because they had not been studied before. It seemed reasonable to suppose that, for example, different target types would be susceptible to different probabilities of identification under conditions of varied image scale or film quality. If such relationships did exist, the information would be of apparent practical significance and might stimulate additional investigations to improve military interpretation system capabilities.

Because of the fairly large number of independent variables which could be studied, it was of first importance that they be grouped in terms of their functional aspects. Such grouping not only lends orderliness to the investigations, but permits more complete appreciation of interrelationships of study findings.

It was believed that the several measures could be convenently conceptualized in terms of whether they might be considered as non-controllable or controllable variables. The non-controllable variables are considered to be such factors as the types of targets which might be present within a specified image, the number of such targets present within a single image (image complexity), the physical position of the targets within the image, and the figure-ground contrast of the targets (target conspicuity). The controllable variables include such image factors as film density and scale, and such display variables as mode, presentation rate and display size.

If these variables are discovered to have a measurable influence upon interpreter performance, changes may be indicated for the method of data collection or data display (for controllable variables) or changes may be made in operational requirements and personnel assignments to attempt to meet the problems imposed by the non-controllable factors.

The implications of the various factors will be discussed in the next section. This section is concerned with defining the variables, describing the methods which were used to measure each, and presenting the results of analytic efforts relating each measure to various of the criterion measures.

B. DEFINITION AND MEASUREMENT

1. Non-Controllable Variables. Except during the rare circumstance of photographing "targets" for specific sorts of interpretation research, the image data collection agency has little or no control over certain features of the resulting photography.

The agency does not control, for example, the type of targets which may be grouped in an area, nor does it control their quantity, their deployment, or their "visibility." These are transient characteristics under the control (in a military situation) of the opposing forces. The detection of such features and the detection of their changes is the primary purpose of the imagery interpretation efforts.

Although the features noted above are not directly controllable, it is important to understand how their variations may influence image interpretation efforts. Then, depending upon the nature of the interpretation task, allowances may be made in order to improve interpretation accuracy. These allowances may include, for example, a relaxation in time requirements, or the use of interpretation teams, or the employment of electronic aids to "enhance" certain types of target outlines.

The point is that knowledge of relationships between noncontrollable variables and interpreter performance provides information upon which to base specific techniques in order to overcome deficiencies in well-defined situations. It is a specific, rather than a "shotgun," approach to the solution of practical problems.

For present study purposes, three noncontrollable variables were selected for study: target type, number of targets per image ("complexity") and target-to-background contrast ("conspicuity"). Each of the analyses involving these three variables is described below.

a. Target Type. The subjects' Completeness scores (combined Experimental and Control groups) were grouped in terms of the six major target family types in accordance with the image scale at which the targets were displayed. The purpose of this sort of grouping was to ascertain if image scale was an important factor in determining whether specifiable target types would be identified.

In terms of image scale, there appeared to be little difference for large or medium scale imagery (Completeness scores of .62 and .68, respectively) but Completeness scores for small scale imagery dropped to a median value of .48.

Of the six target families, only Airfields and Port/Harbor appeared to be impervious to small scale effects, maintaining essentially similar Completeness scores across the scale categories. All other target families indicated a slight rise in Completeness scores for medium scale imagery and a marked drop for small scale imagery.

The median Completeness scores for the combined Experimental and Control group subjects (Projected Display Test, all image scale categories) are presented below:

Target Family	Completeness
Military	68%
Airfields	67%
Power	66%
Rail	61%
Industry	54%
Port/Harbor	39%

b. Target-Background Contrast. Another variable over which the image collecting agency has no direct control concerns the visibility of a target against its surrounds. For present study purposes, this feature is defined as "conspicuity."

The conspicuity of each target was rated independently by two judges using a three-point rating scale. Targets which contrasted highly with their background received a rating of one and those which blended with the background received the three-point rating. Interjudge agreement using this scale was 83 per cent, and disagreements were resolved in discussion sessions.

There were two basic aims in testing for possible conspicuity influences upon performance. First, there was an interest in determining whether ratings of image quality variables could be reliably made.

It seemed fairly evident that if judges found that certain targets were difficult to see under unlimited time conditions, the test subjects under conditions of rapid viewing would have at least equal difficulty. Consequently, a test of this point would appear to be confirming the obvious. However, if reliable ratings of target conspicuity could be simply derived, and if these ratings did in fact relate to later performance measures, then this information could be usefully applied. The over-all worth of a roll of aerial imagery could be assessed in terms of the probable level of interpretation completeness which could be achieved. The assessment would consist simply of subjectively rating the imagery in terms of a few quality dimensions, of which conspicuity is one.

A second aim of this particular analysis was to provide some basic information concerning possible interrelationships among target type, conspicuity ratings, and Completeness scores. These data were considered desirable because of the increasing attention being given to image interpretation. A fairly large number of research teams are performing studies, and it is likely that their findings will in few instances be mutually relatable. Unless the criterion test materials for each study cover a broad range of target families and image qualities, it is likely that biases will be introduced in test results. If the natures of these possible biases are understood, they can perhaps be avoided.

The total subjects' Completeness scores on the Projected Display imagery were calculated in terms of conspicuity ratings within each of the image scale categories; and the total subjects' Completeness scores on the Light Table Test were calculated in terms of conspicuity ratings but without regard to image scale. These data are presented in Figure 2-1. It will be noted that there is a sharp decrease in mean Completeness scores across the conspicuity rating categories. The single exception for large scale imagery at the figure/ground conspicuity rating of one is probably attributable to random fluctuation because of the small sample size for this category.

Figure 2-2 presents a similar comparison of Completeness scores and target conspicuity ratings but this time in terms of target family categories. It will be noted that poor figure/ground contrast has little effect upon the completeness of identification for Industrial targets or for Airfields. The effect is severe, however, for Power and for Military facilities.

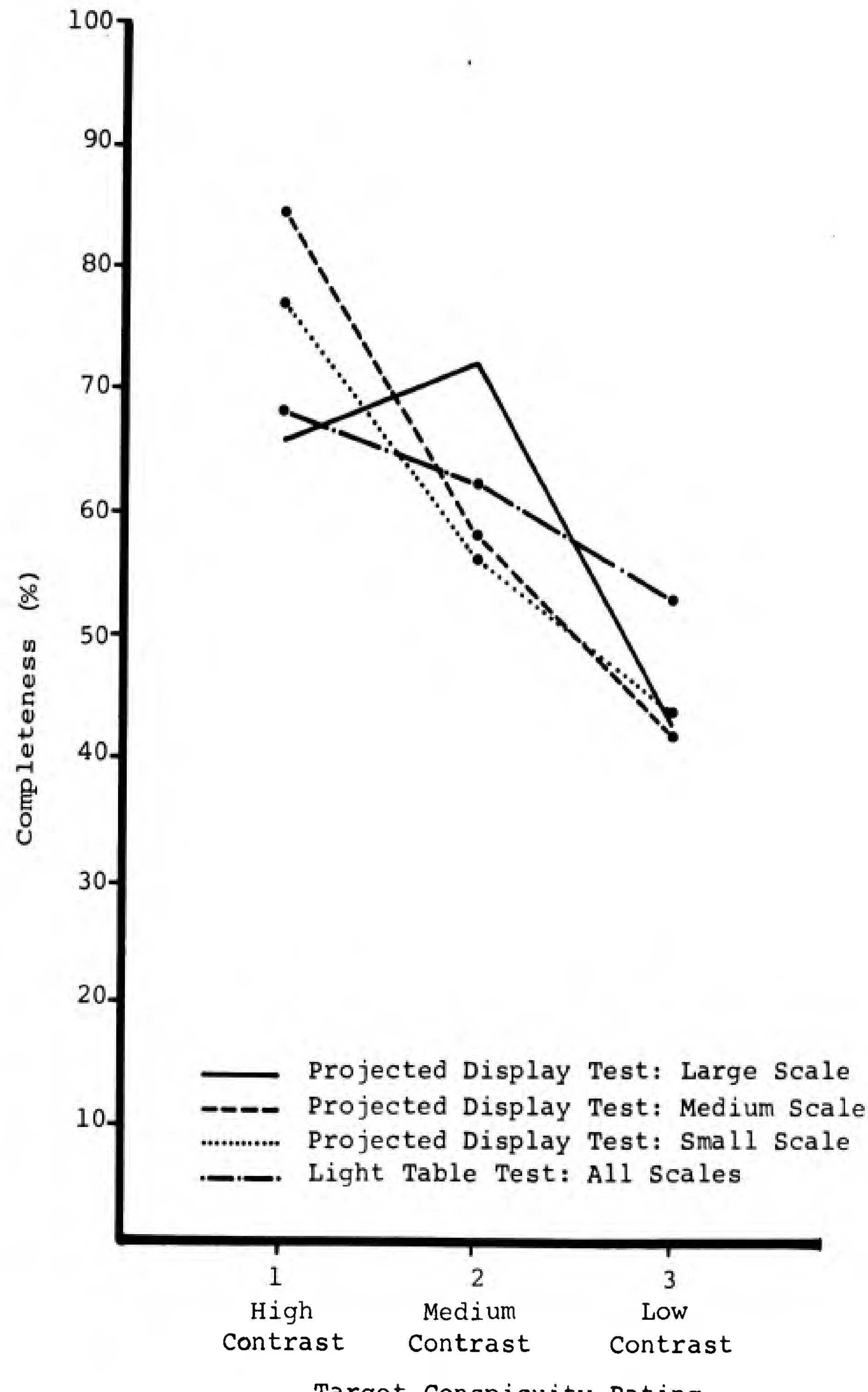


FIGURE 2-1. Influence of Target Conspicuity Upon Combined Completeness Scores of all Subjects for Two Display Modes.

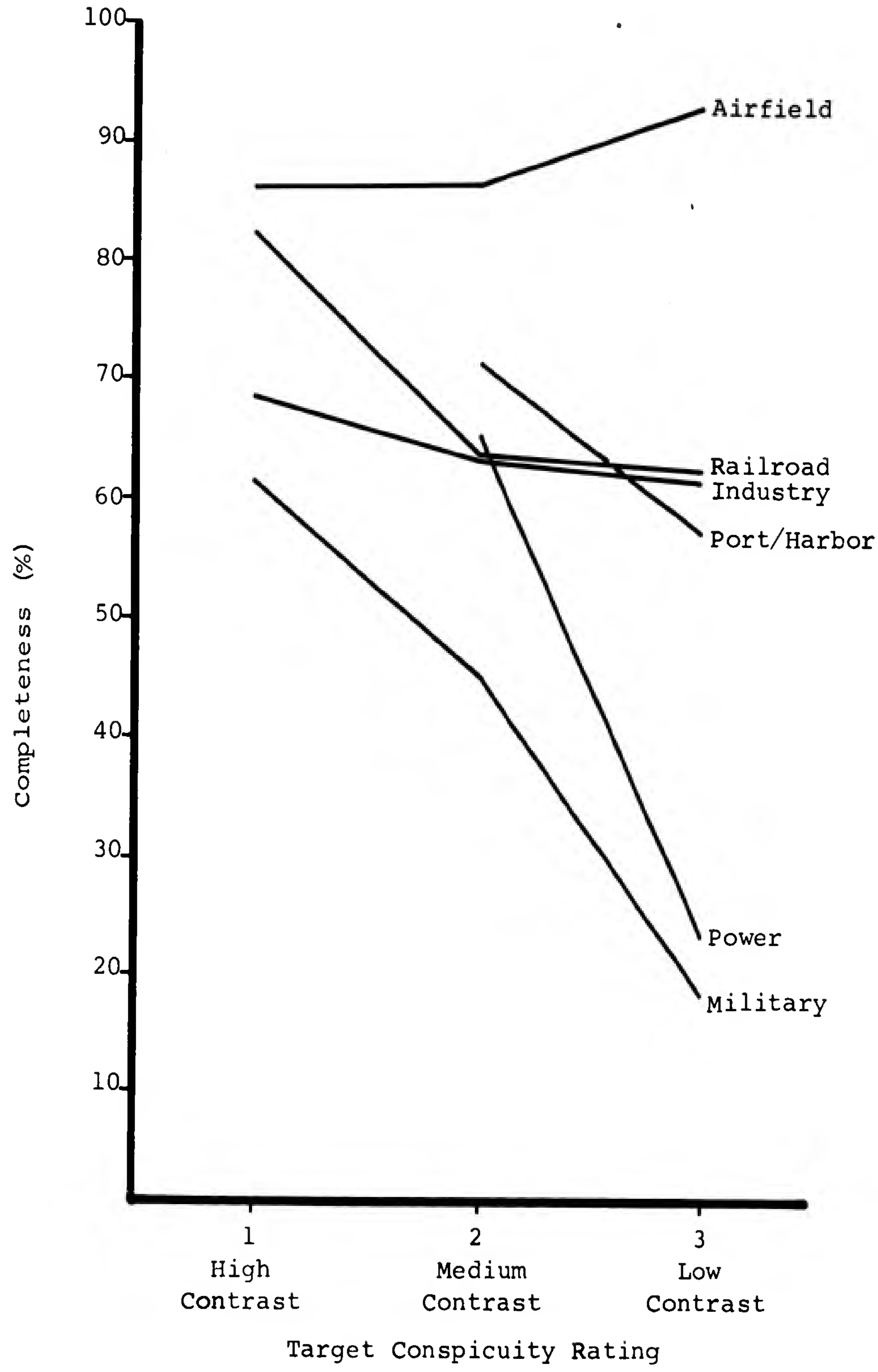


FIGURE 2-2. Influence of Target Conspicuity Upon Light Table Test Completeness Scores in Terms of Major Target Types (Experimental and Control Group Scores Combined).

c. Image Complexity. The influence of image complexity* upon interpreter efficiency is graphically portrayed in Figure 2-3. This graph also illustrates the nature of certain of the "improvements" instilled in the Experimental group during the training program.

For this analysis, Completeness and viewing time values for the Experimental and Control groups were calculated separately in terms of images which had been sorted according to the number of different types of targets they contained. The "Effectiveness" scores were obtained by dividing Completeness by Time, and were converted to "Effectiveness Indexes" on the basis of setting the Experimental group single target score equal to 100.

With respect to training course value, it can be seen that the Experimental group was identifying targets within the more complex groups (four and five targets per image) at an effectiveness level achieved by the Control group only for the least complex (single target) imagery.

The close shape correspondence between the Experimental and Control group curves indicates that approximately twice as much time is required to achieve a specified level of Completeness when viewing "complex" imagery as when viewing single-target imagery.

In addition to the over-all Completeness/Time (Effectiveness) Index, separate plots were made of the distribution of Completeness and Time scores for the two groups. Figure 2-4 presents the mean Completeness scores achieved by each group for each level of image complexity. The consistent

---

*Complexity in this context is defined in terms of the number of different target family categories appearing within a single display.

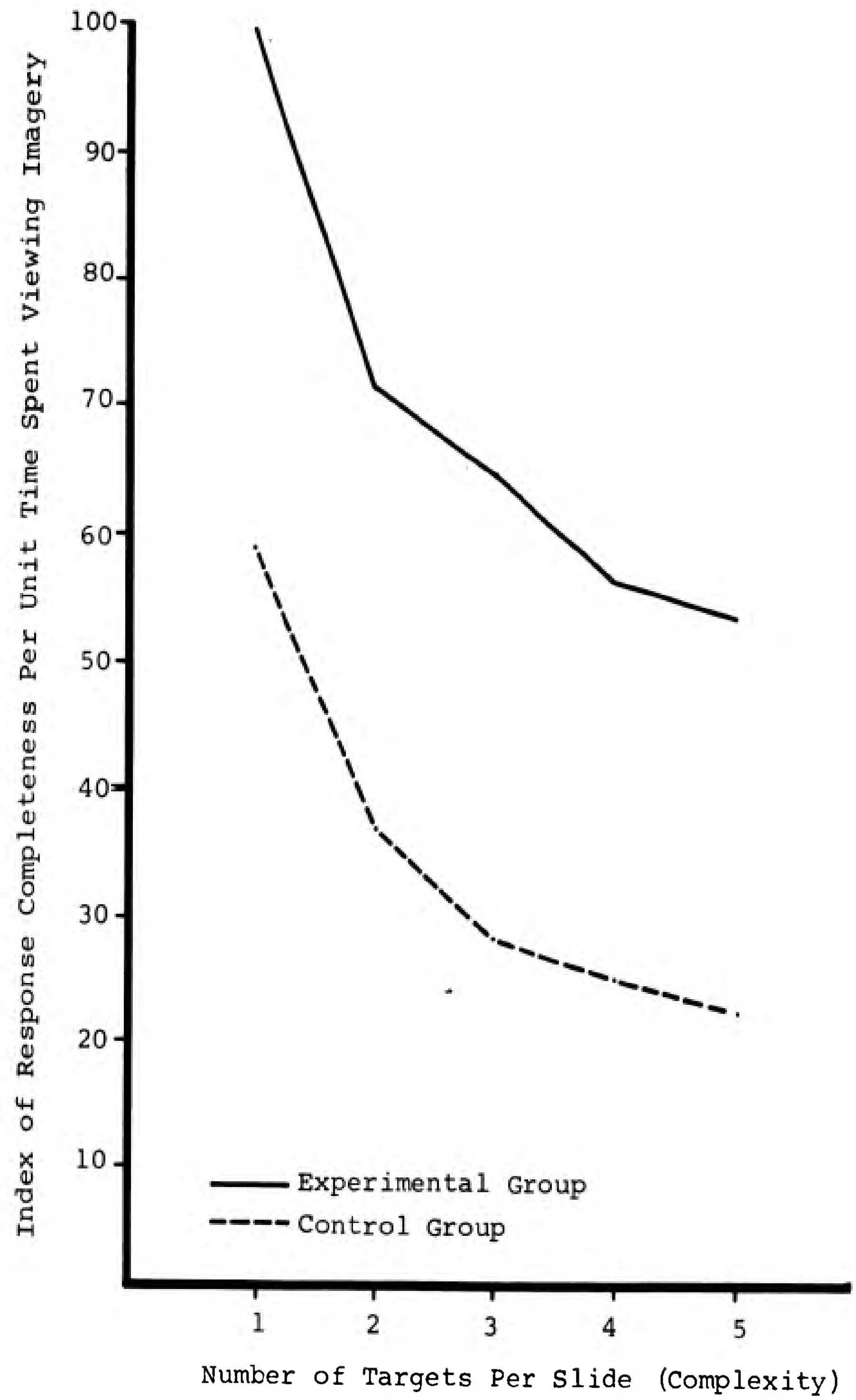


FIGURE 2-3. Comparison of Experimental and Control Group Effectiveness Indexes for Imagery of Varying Complexity (Projected Display Test).

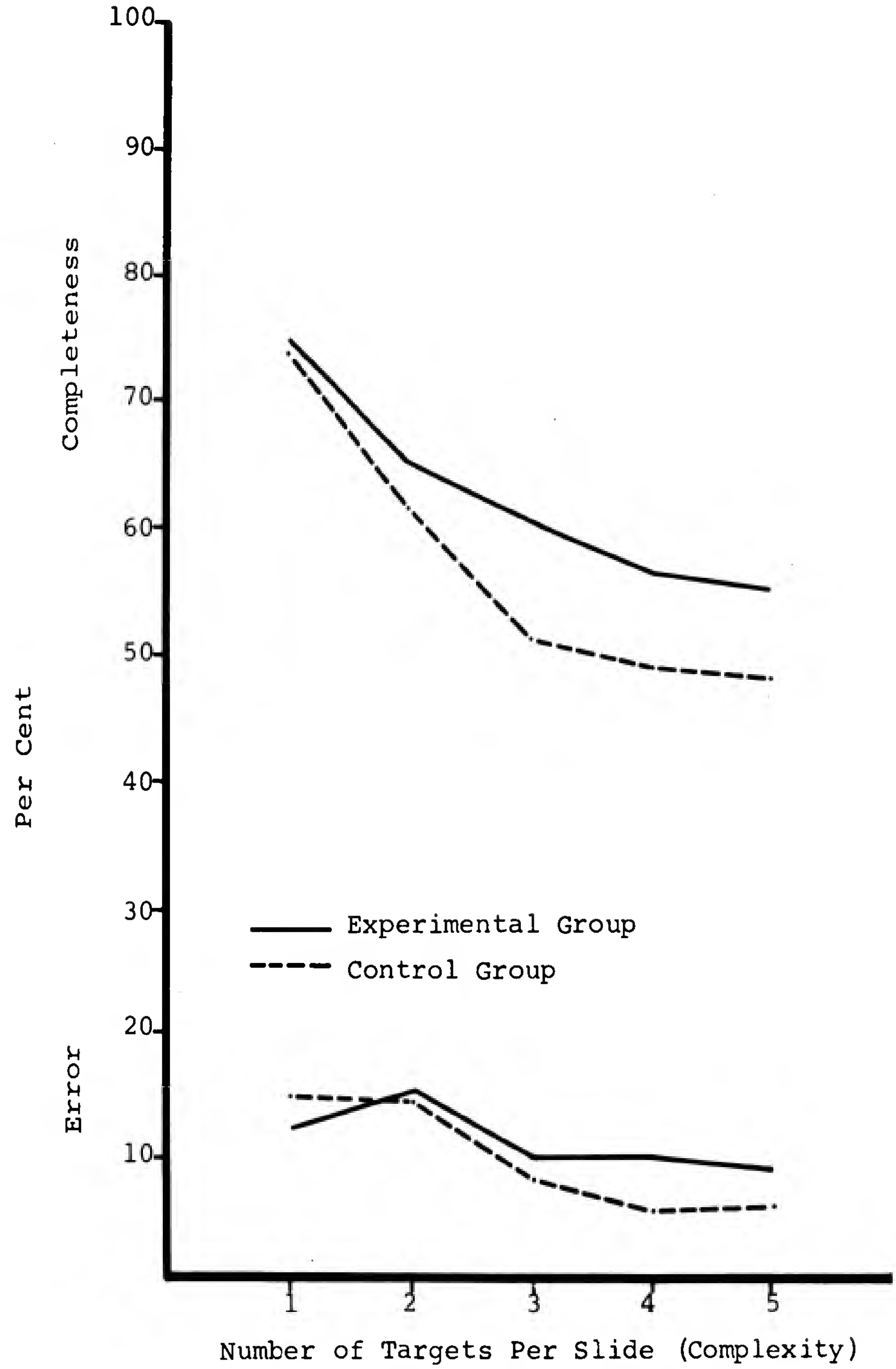


FIGURE 2-4. Comparison of Experimental and Control Group Completeness and Error Scores for Imagery of Varying Complexity.

superiority of the trained group is apparent. In addition, however, plots of error rate are included. There is an indication that the Experimental group provides a slightly higher percentage of incorrect responses for more complex imagery than does the Control group.

This error tendency is reflected in the Accuracy scores achieved by the two groups (see Figure 2-5). Since this Accuracy score is computed as the percentage of total responses which are correct, the error scores of the trained group tend to reduce their over-all accuracy for four and five target per slide imagery.

Considered together, Figures 2-4 and 2-5 indicate that the trained group tends to provide a larger number of target identifications in a much shorter time than does the Control group. This productivity raises the Completeness score but, since not all identifications are correct, the Experimental group accuracy is impaired.

A plot of the mean time spent by the Experimental and Control group members in viewing each image in the Projected Display Test is presented in Figure 2-6. However, in this graph, the images are sorted in terms of scale as well as complexity. The figure graphically illustrates that training, scale, and complexity interact in their influences upon viewing time. One effect of the training program appeared to be a sharp reduction in the amount of time required to identify targets regardless of scale and with only a relatively slight influence from complexity. The Control group evidences the substantial nature of the interacting effects of the Image Scale and Complexity variables in the absence of rapid recognition and target signature training.

2. Controllable Variables. Among the variables under at

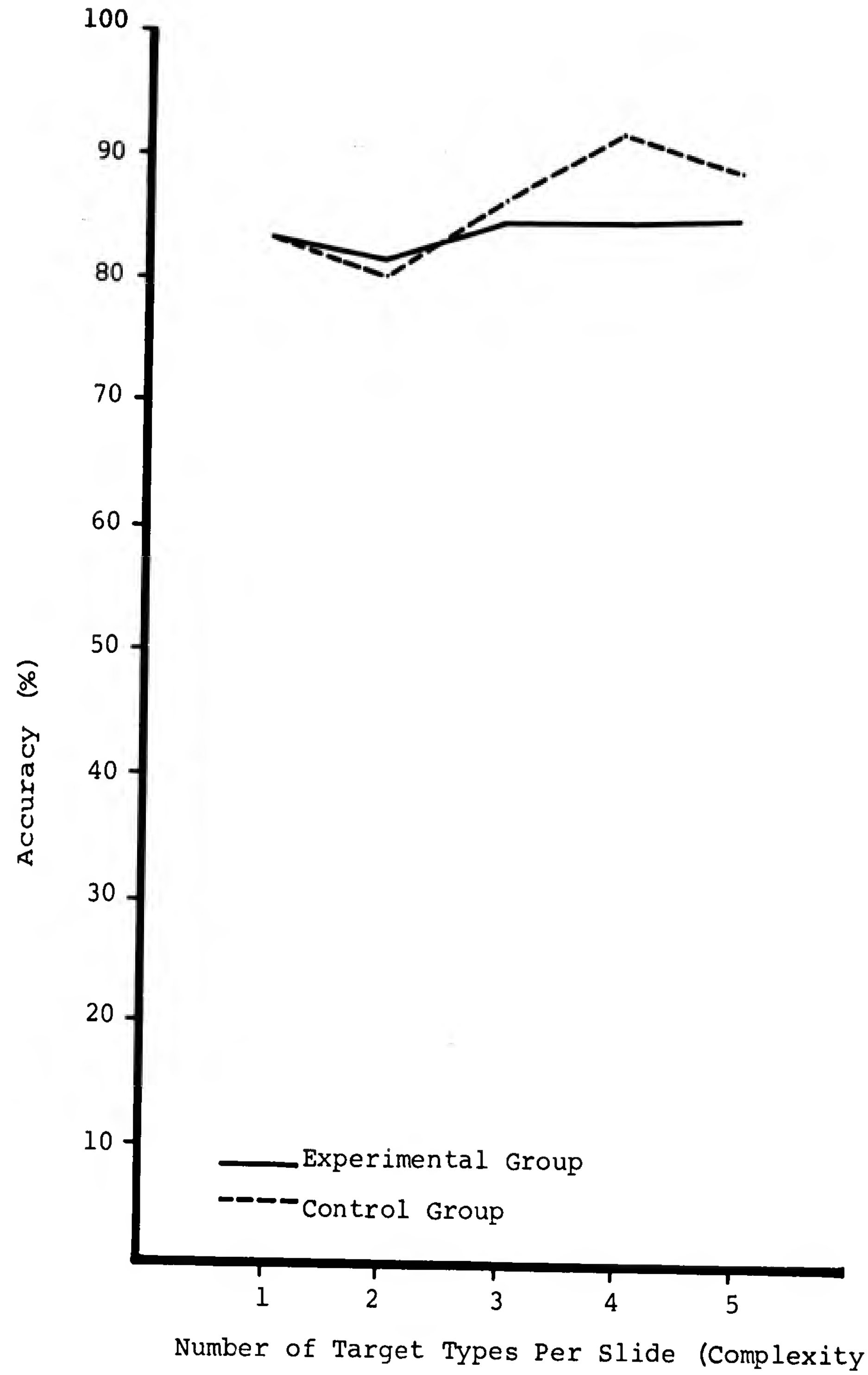


FIGURE 2-5. Comparison of Experimental and Control Group Accuracy Scores for Imagery of Varying Complexity.

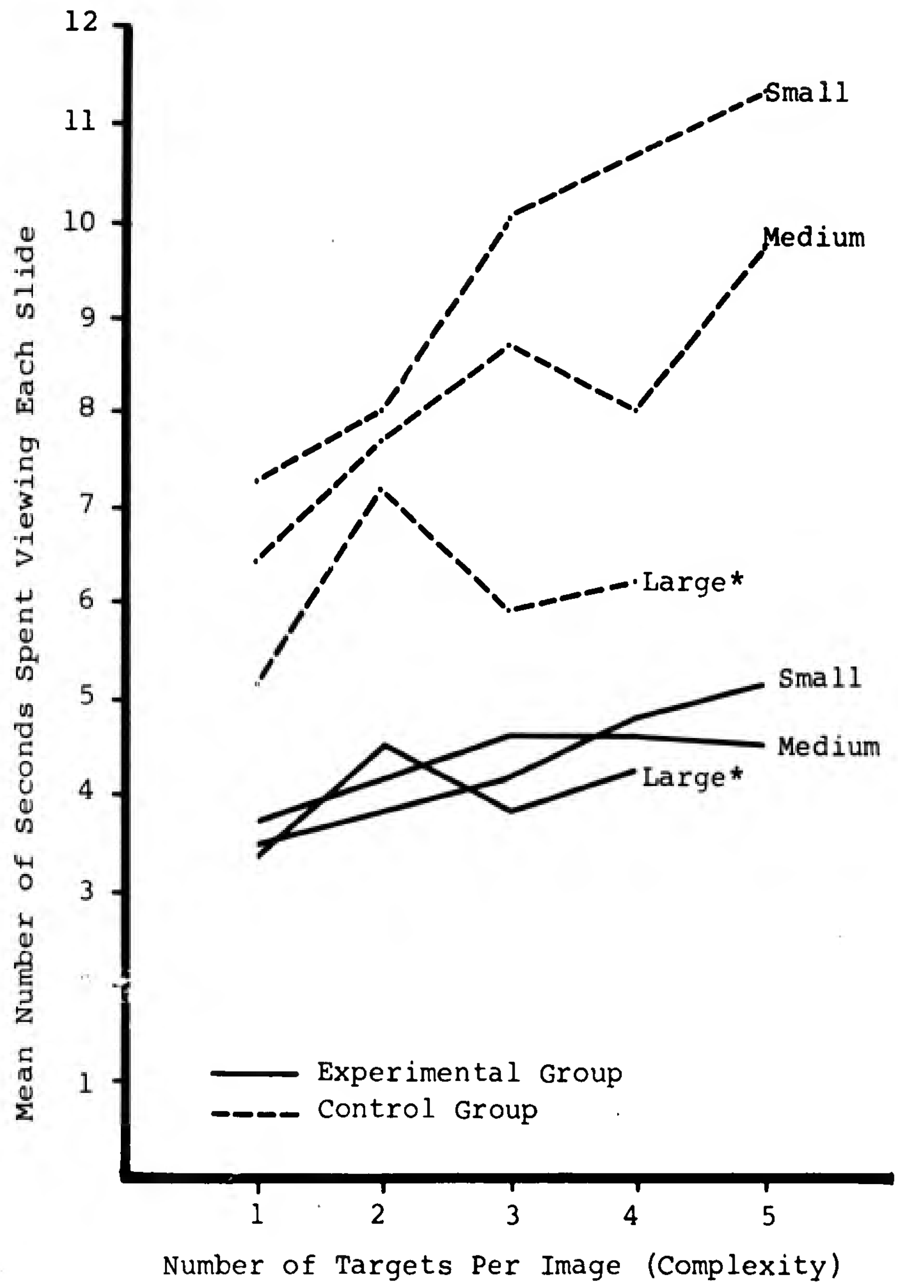


FIGURE 2-6. Comparison of Experimental and Control Groups in Terms of Mean Time Spent Per Image as a Function of Image Scale and Complexity - Projected Display Test.

*No Large Scale Imagery was available which contained more than four different targets.

least partial control of the image interpretation agency are such features as image scale, film density, target size, and the mode of displaying the processed films. Of these, the variable of scale has been included in a majority of the analyses thus far reported and need not be repeated here. In general, the results indicate that the image scale has a decided influence upon all criterion measures with, as might be anticipated, large scale imagery generally permitting the highest levels of achievement.

The variable of film density is under control of the data collecting agency in the sense that density is governed by the type of film employed, the exposure settings used, and the chemicals and procedures used in developing the exposed film. In the present study, ratings were made of film density in the same manner as that employed in rating target conspicuity. However, upon tabulation of the ratings it was found that only four images had been rated as "dense," perhaps because of earlier efforts in weeding out low quality imagery. Consequently, despite a high percentage of agreement on the remaining ratings, no further analyses were conducted for the density variable.

A third controllable variable concerns the mode of display employed in viewing the processed film. Customarily, a light table has been employed for this purpose; the film is either manually or electrically transported across a ground glass beneath which lights have been installed.

Recently, increased attention is being given to rear projection displays. These have a number of advantages, including a magnification of the display, convenience in film handling, etc. However, they are not in widespread use and at least some interpreters feel that the rear projection screen introduces an undesirable degree of resolution loss.

Because of the likelihood that rear projection displays will be increasingly used, it appeared that some data concerning relative performance under the two display modes would be helpful.

a. Display Mode. The 17 three-frame sequences included in the Light Table Test materials contained six frames, portions of which had been reproduced in slide format for inclusion in the Projected Display Test materials. The six duplicate portions contained 18 targets in common, worth a total of 32 "points." Although it is somewhat artificial to regard these targets as a "subtest," the common portions of the two tests appeared to provide data concerning relative performance which could be expected under two different display conditions.

The comparison of Experimental and Control group performance for the two display modes was based upon several assumptions. First, it is appropriate to note that the Light Table Test was administered to each subject after the Projected Display Test had been administered. Consequently, subject performance on the duplicate frames might be expected to benefit from the previous viewing. It was assumed that such recall effects would be minimal in the present instance because of the large number (215) of Projected Display Test targets to which the subjects had been exposed. Also, since no feedback had been provided concerning the accuracy of the subjects' identifications on the Projected Display Test, there was little a priori support for surmising that their responses would necessarily improve appreciably during a second viewing.

The comparison of performance, since it is partially based upon time, requires additional assumptions concerning individual target viewing times and response recording times.

Only the total test time had been recorded for the Light Table Test, but times spent viewing individual images had been recorded for the Projected Display Test. Since the Display Mode comparison was to be based upon identical portions of the test materials, it was necessary to convert time values to a common base, e.g., mean time required per target identification. The two groups' total times for the Projected Display Test were divided by the total number of targets (215), and these mean time-per-target values were compared with the actual times which had been recorded during administration of the test. The close correspondence (within 3 per cent) of the estimated and the measured "time-per-target" provided support for using estimates of time-per-target values for the Light Table Test.

However, the Light Table Test time-per-target value included the amount of time spent by the subject recording his identification. The Projected Display Test time value included only the time that each image was actually being visually searched. Therefore, unless a correction was applied to the Light Table Test time, the two values would not be directly comparable. The final assumption, consequently, and one which was based upon surreptitious measures of several subjects while they were taking the test, was that an approximate mean time of four seconds was required to locate the appropriate line on the answer sheet and write the abbreviated name for an identified target. For each group, the total number of responses made on the Light Table Test was multiplied by four seconds to derive the estimate of total time spent in recording answers. This value was subtracted from each group's total time score and the resultant figure was used in the further comparisons of group performance.

The first analysis indicated that the assumption concerning an absence of practice effect was ill-founded. As indicated in Table 2-1, the groups' Completeness scores (total minus "subtest") are fairly comparable regardless of Display Mode. However, the Completeness scores obtained on the "subtest" were substantially higher on the second administration of the frames.

TABLE 2-1

MEAN COMPLETENESS SCORES FOR TWO DISPLAY MODES

	Projected Display	Light Table Display
Experimental Group		
"Subtest"	60%	85%
Total minus "Subtest"	62%	60%
Control Group		
"Subtest"	58%	77%
Total minus "Subtest"	56%	55%

Despite the obvious improvement introduced in the Test by the repetition of six images which had previously been displayed, the differences between the display modes were still substantially in favor of the Projected Display. The differences favoring rear projection viewing were discovered for those measures involving time. On the average, it required seven to ten times longer to achieve a comparable level of Completeness when using the Light Table Display than when using the Projected Display.* This summary statement is based upon results presented

*It must be recognized that the display modes differed also in terms of the duration that imagery was exposed to view. Although the subject could call up repeated views of an image on the Projected Display test, each demand provided only a two-second view. This condition did not apply for the Light Table test. Consequently, differences between the two modes may be a function of the display or they may be a function of the machine-paced exposure durations.

in Table 2-2. For each display mode, the reader's attention is also directed to the difference between Experimental and Control group scores. Both groups perform much more rapidly on the Projected Display but the Experimental group still requires less than half the time required by Control group members.

TABLE 2-2

MEAN NUMBER OF SECONDS REQUIRED TO IDENTIFY EACH TARGET UNDER TWO DISPLAY MODES

	Projected Display	Light Table Display
Experimental Group		
"Subtest"	3.31	23.09
Total minus "Subtest"	2.74	26.09
Control Group		
"Subtest"	6.76	35.71
Total minus "Subtest"	5.73	39.78

This point is graphically presented in Figure 2-7, which combines the Completeness and Time measures into an "effectiveness" index, or the relative percentage of completeness achieved during comparable time periods when using the two display modes.

b. Target Size. Although the actual physical size of targets may be considered a noncontrollable variable in the sense that this term is used in the present report, the size of the target on the display is controllable. Target size on a display may be partially controlled before the fact by selecting a desired combination of flight altitude and lens focal length, or it may be controlled after the fact by magnifying portions of the display.

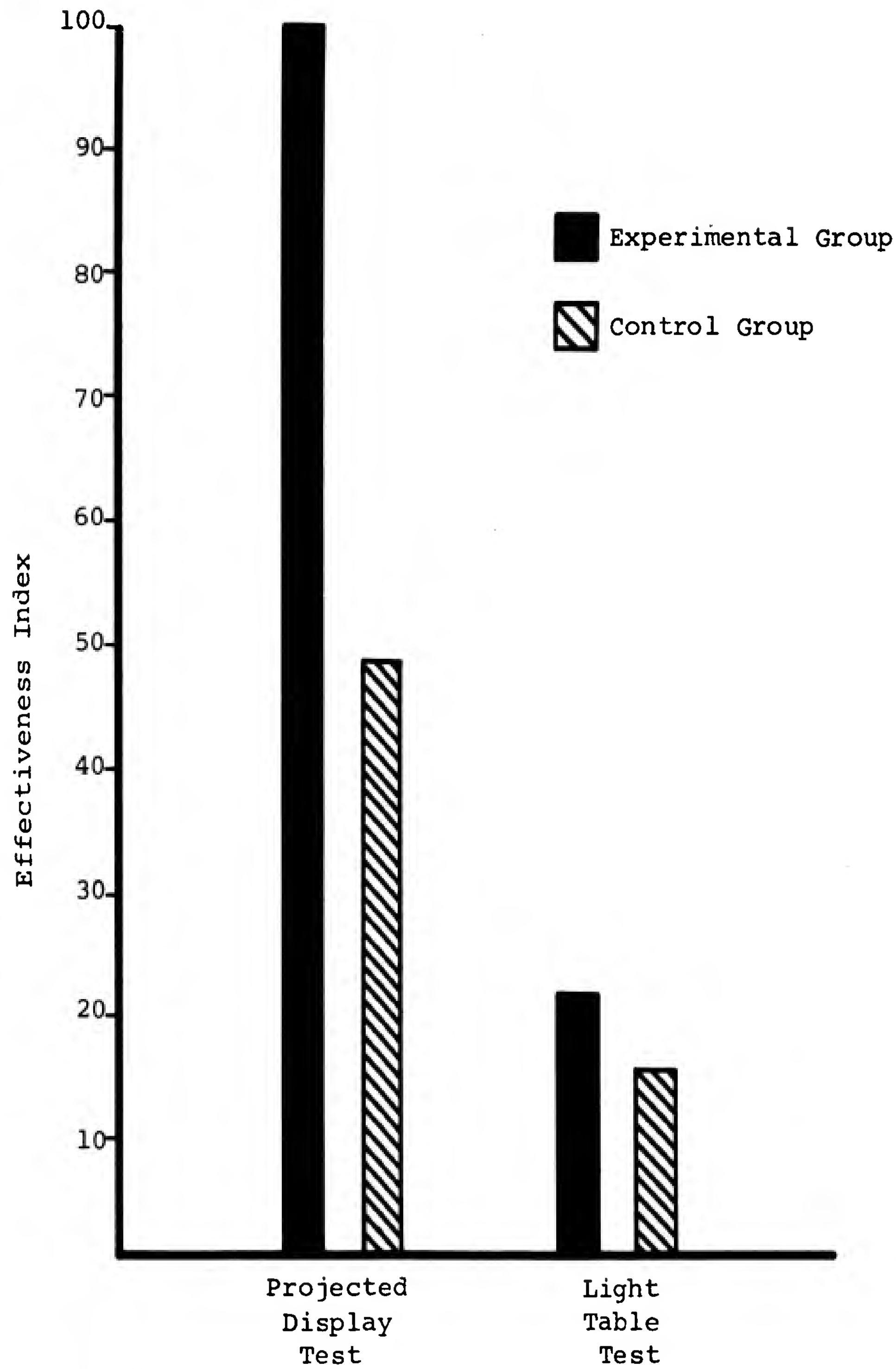


FIGURE 2-7. Comparison of Two Display Modes in Terms of Experimental and Control Group Effectiveness Indexes.*

*Completeness Score divided by Time

One primary purpose of the present investigation was to determine whether magnification was "necessary" in the type of image screening task under study. The projected display was, by its nature, a 2.2-power magnification of the original imagery and the target family categories were, for the most part, physically quite large. Consequently, it was felt that, with rare exceptions, performance scores could be expected to be fairly uniform across all target sizes save the very small, and that these "identification curves" could be expected to vary with target type.

If these assumptions were found to be correct, the "identification curve" plots would have immediate practical utility. The rationale for this belief is as follows.

Since a definite relationship would appear to exist between the level of detail it is possible to extract from an image and the size of the target of interest within that image, and, since different target families would appear to have different size requirements in order for an interpreter to achieve a given level of completeness of identification, it would appear that useful guidelines are potentially available for image collection purposes. If the approximate physical dimensions of a target type are known, then the image size of that target can be calculated for any selected image scale. Therefore, the minimum necessary scale value ("critical scale value") can be established, depending upon the level of interpretation desired; and missions to collect imagery of certain targets can be planned in terms of collecting imagery at least as large as this critical scale value.

In order to test the rationale and its related assumptions, two measures of target size were obtained. The first was simply a measure of the longest dimension (in inches, or decimal portions thereof) of the portrayed target. For fairly compact targets, such as antennas or power plants, this was a practical and meaningful measure. For elongated targets such as rail lines and harbor areas, however, the measure left something to be desired. Consequently, a second measure was calculated; that of target area. This was expressed as the percentage of the display area which was encompassed by the target area.

The first analysis which was conducted compared Experimental and Control group performance on the Projected Display Test for all targets. The criterion was the number of correct "points" obtained by each group in terms of varying target sizes. The results of this analysis are graphically portrayed in Figure 2-8.

The analysis revealed that, although the Experimental group consistently scored higher than did the Control group, the two curves are nearly identical in shape. The reader will recall that such group similarity did not exist when analyses were conducted of the effect of image scale. Two conclusions are plausible -- (1) that the interpretation problems relative to target size are (within certain limits) independent of those relative to image scale, and (2) that Experimental and Control group scores may be combined for further target size analysis purposes without fear of biasing the results. All subsequent analyses were, therefore, conducted using the combined group's scores, and nearly all analyses were conducted for the Projected Display Test.*

*The reason for this constraint was that viewing distance was, for this test, held fairly constant at 30 inches. A given target size would, therefore, represent a constant visual angle for all subjects. Analyses were later extended to the Light Table Display for reasons which will become clear. For the present, all references are to the Projected Display Test.

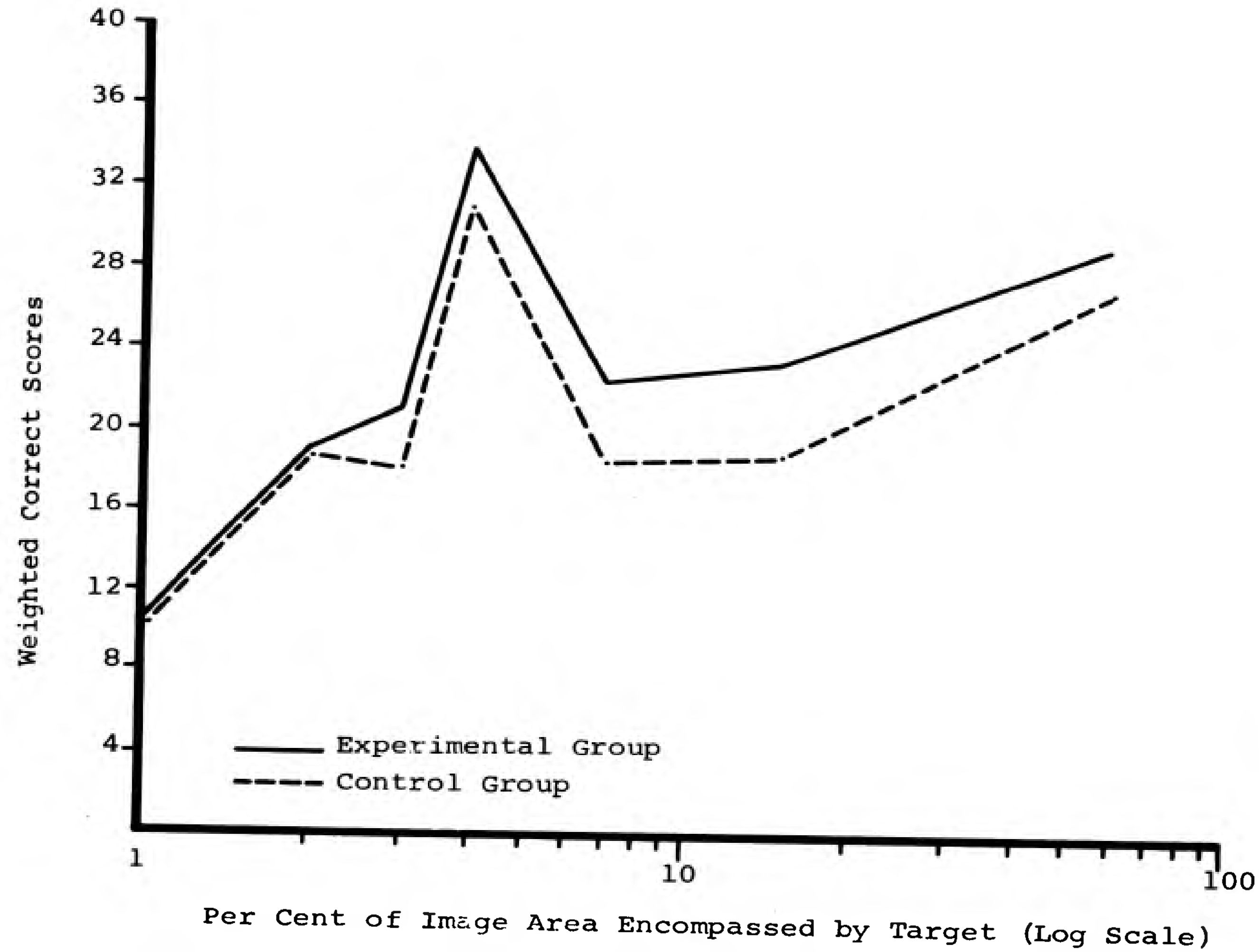


FIGURE 2-8. Comparison of Experimental and Control Group Scores for Targets Encompassing Various Percentages of the Display Area.

Figure 2-9 presents Completeness scores achieved by the combined groups for the six major target types in terms of their longest dimension on the display. The reader will note the generally increasing Completeness scores as target size increases. This trend is more readily apparent in Figure 2-10 which presents smoothed curves for two target families, plotted on a logarithmic scale. In this graph, the Rail category was divided to show rail lines as distinct from marshalling yards, transshipment yards, and roundhouses.

Figure 2-11 presents Completeness curves for Power and Airfield facilities. Once again a linear trend is apparent except that a decreasing trend is noticable for Airfield completeness scores at the larger target sizes.

The Military and Industry categories appear in Figure 2-12 and, for both of these major target families an irregular relationship exists between target size and Completeness of identification by the interpreters. For both types of targets, a definite "dip" in scores is noted after an initial increase in scores. The dip is followed by another increase and another dropoff in scores (similar to that noted for Airfields) at the large target sizes.

Believing that this dip effect may have been due to the grouping within a major target category of a number of targets of drastically unlike configurations, we divided the Military category into two subcategories of "Military Base and Military Supply" and"Military Electronics and Missiles." Completeness scores were replotted and, once again, a marked similarity was found to exist between divergent target configurations (see Figure 2-13). The similarity is apparent although the "Military Base" curve is slightly displaced. The reader

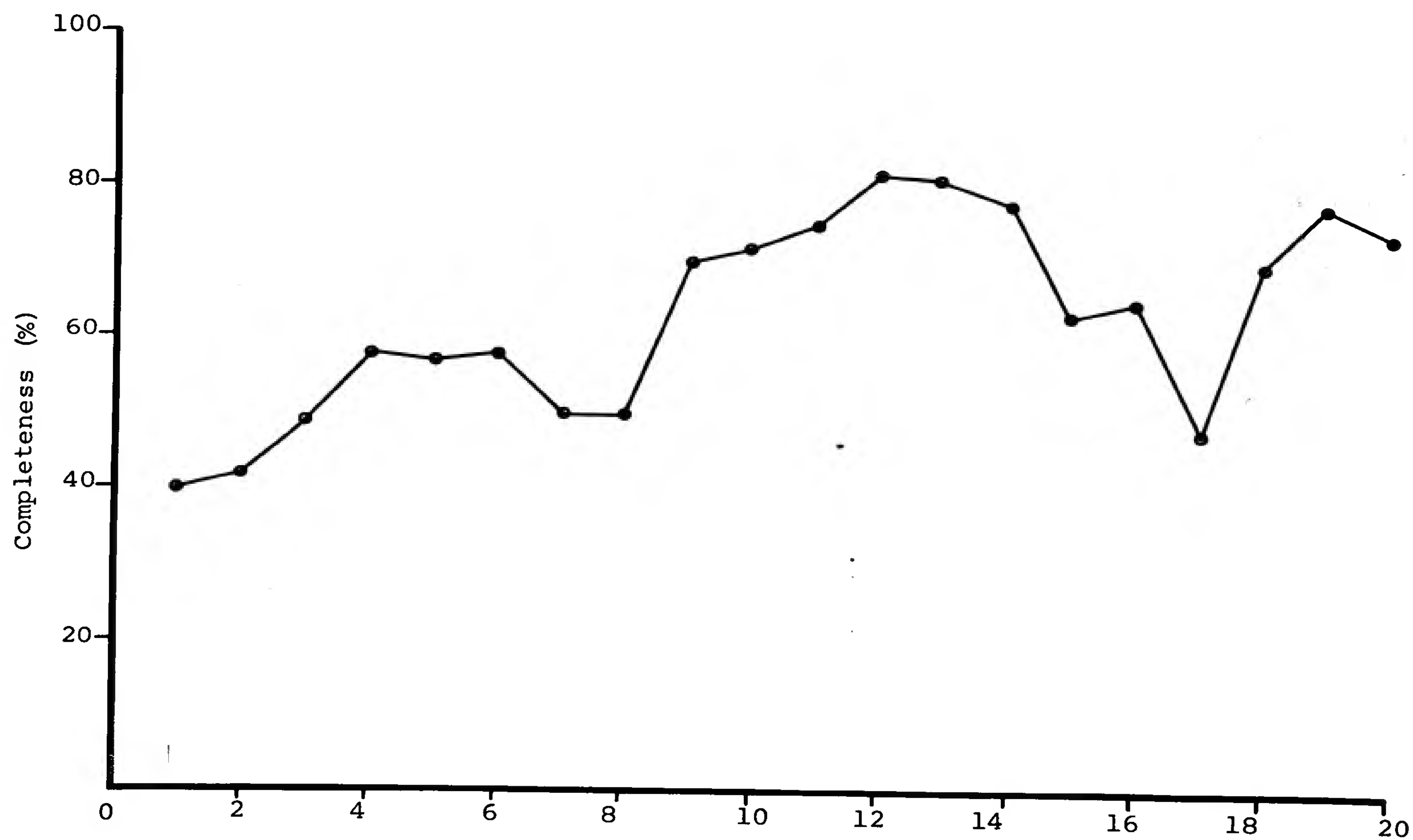


FIGURE 2-9. Target Family Completeness Scores for Targets of Various Size Dimensions. (Experimental and Control Groups Combined, all Target Types Combines)

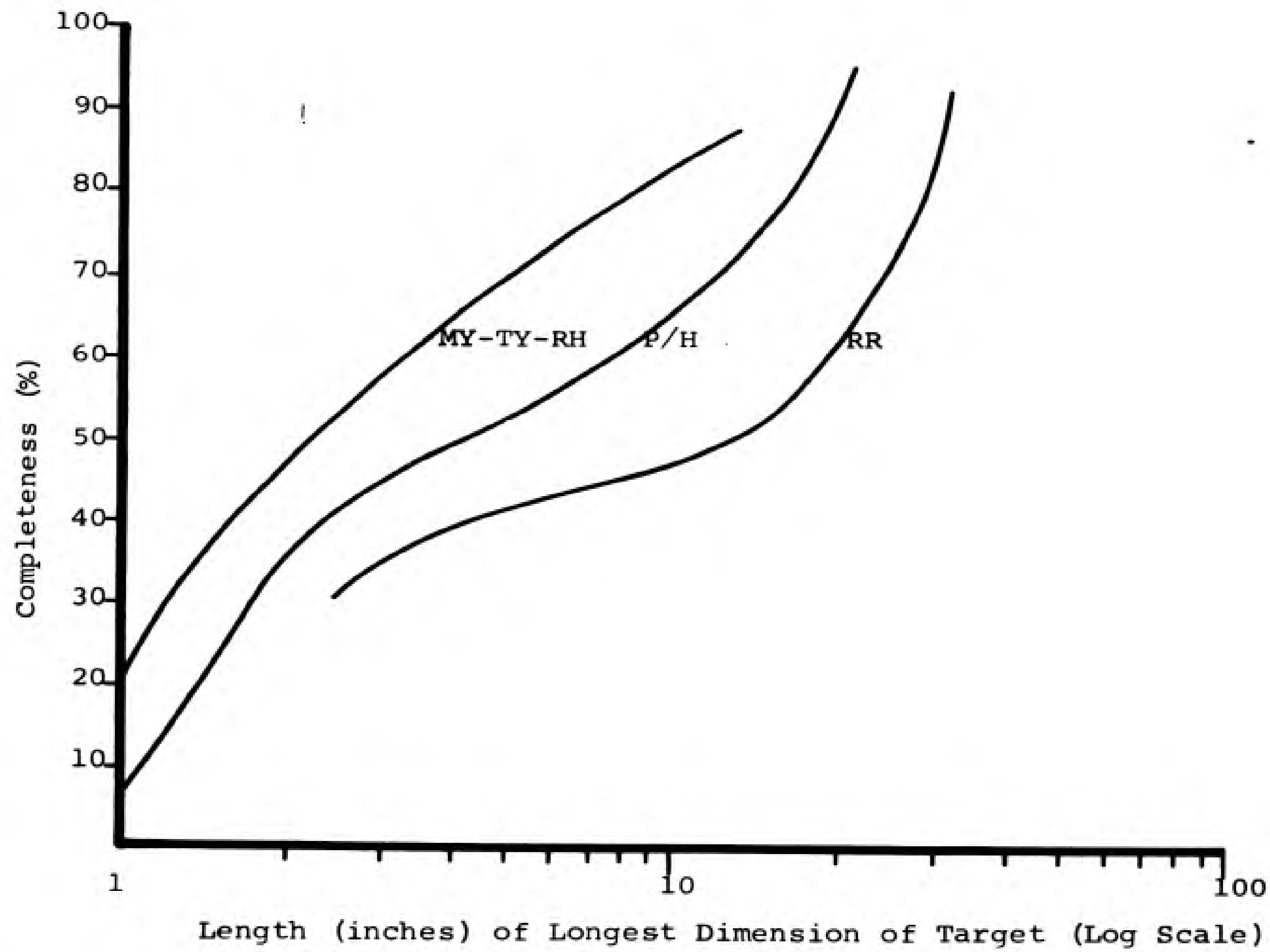


FIGURE 2-10. Smoothed Curves for Selected Target Families. (Experimental and Control Groups Combined)

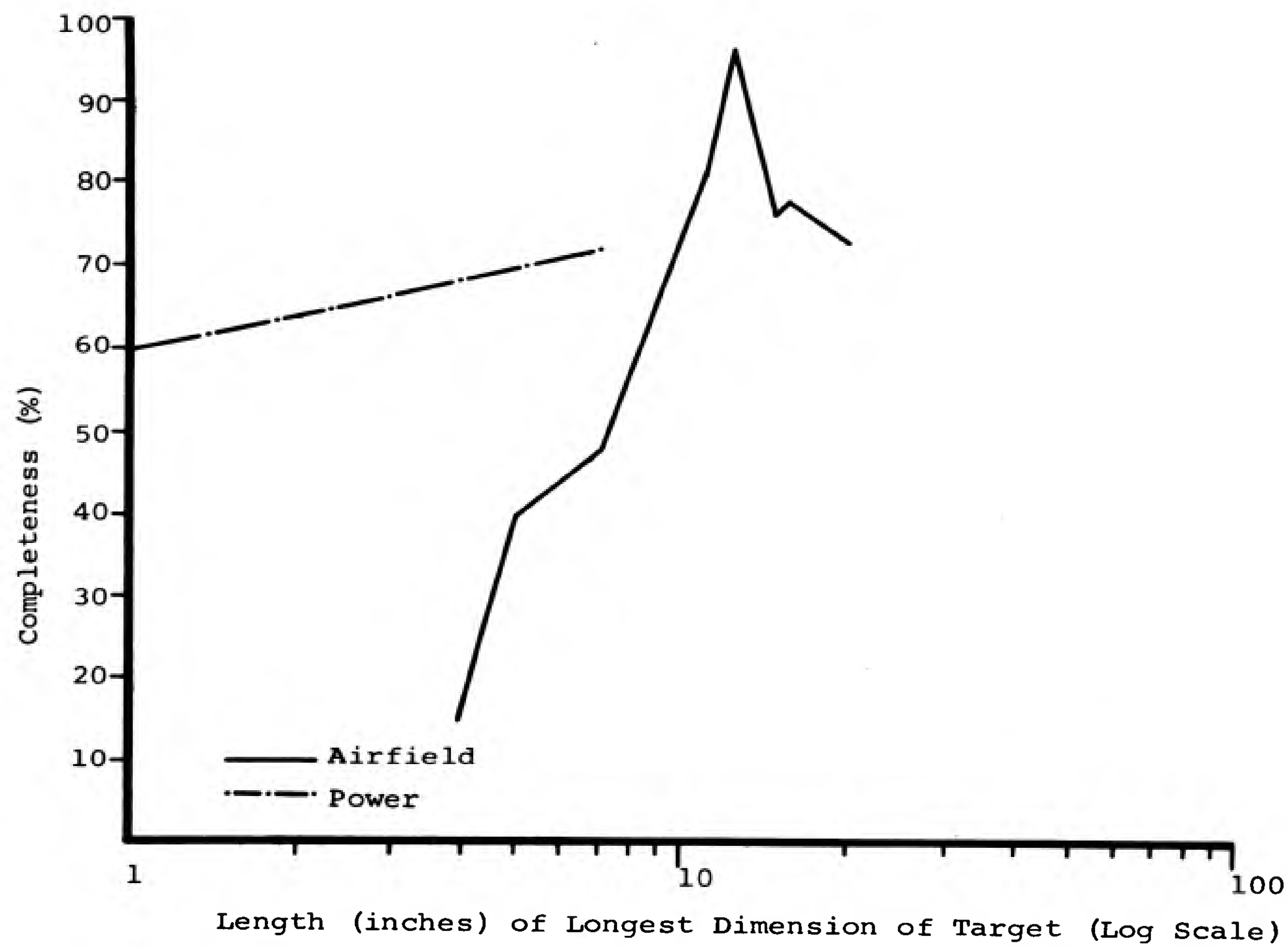


FIGURE 2-11. Power and Airfield Target Completeness Scores for Targets of Various Size Dimensions. (Experimental and Control Groups Combined)

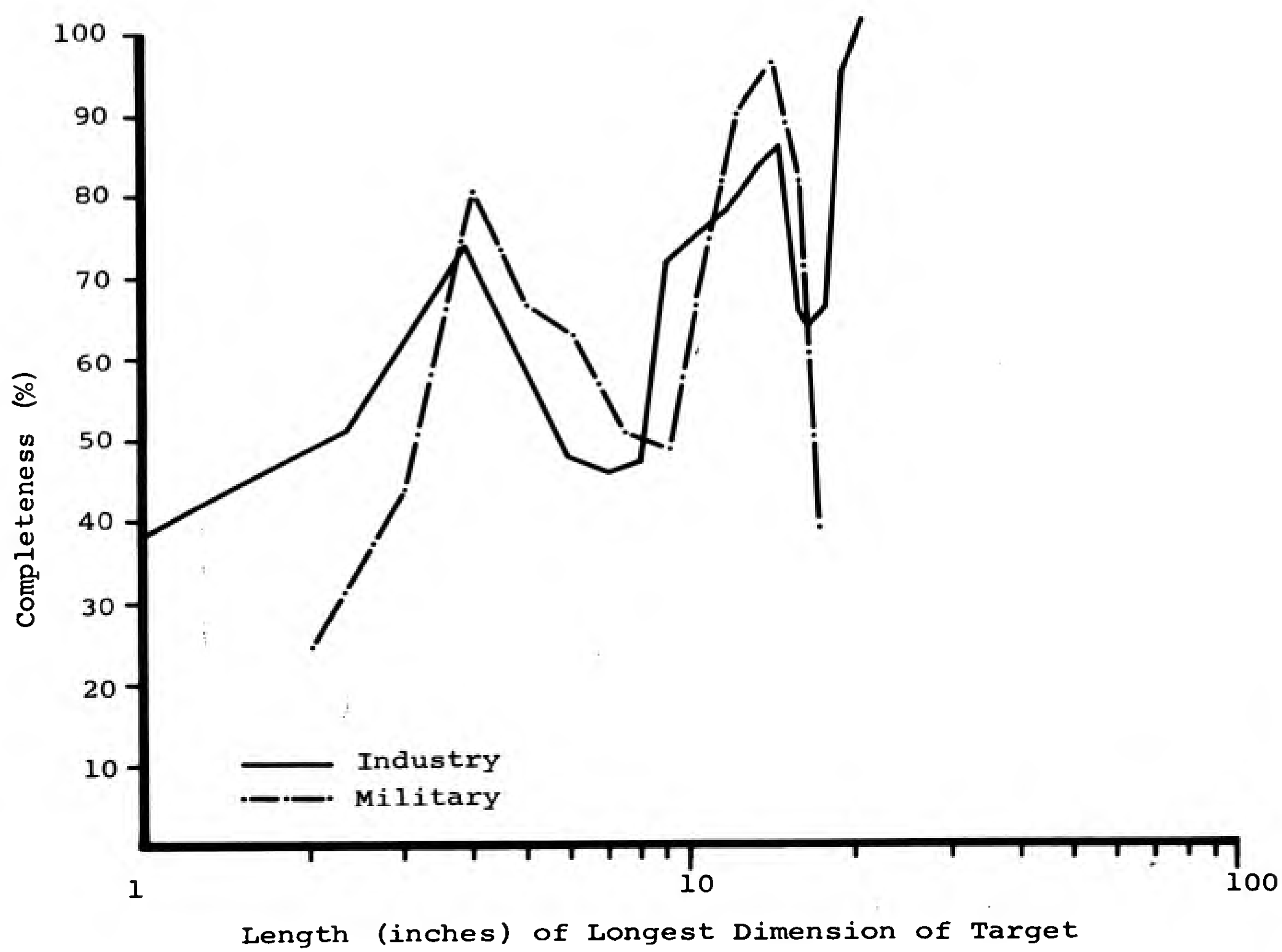


FIGURE 2-12. Industry and Military Target Completeness Scores for Targets of Various Size Dimensions.

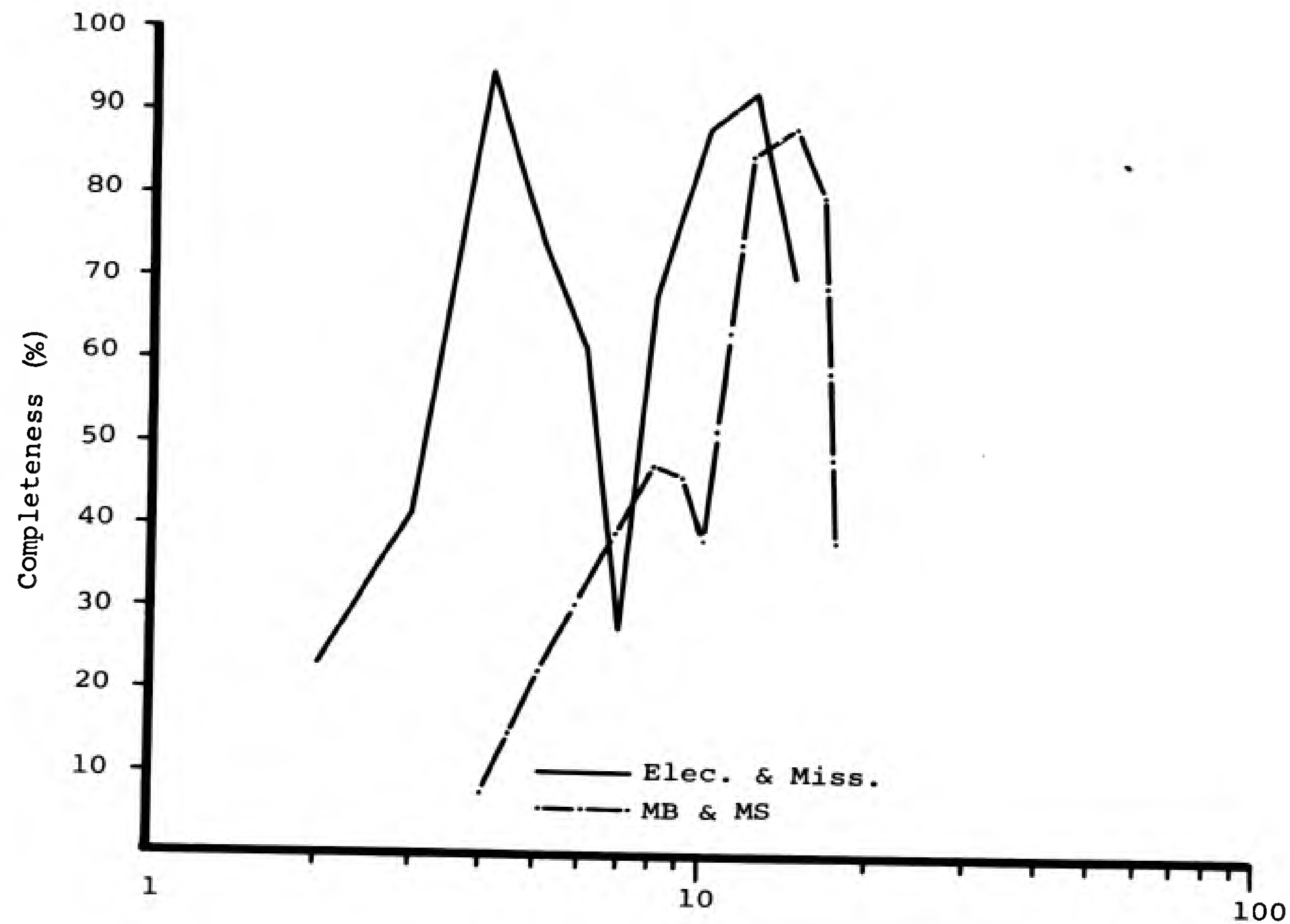


FIGURE 2-13. Two Types of Military Target Completeness Scores for Targets of Various Size Dimensions.

will also note the drop in Completeness scores for the largest target sizes.

The tendency for Completeness scores to <u>decrease</u> as target size <u>increases</u> through mid-range values is an unexpected and puzzling discovery. Several "explanations" which could conceivably account for the finding were considered.

Perhaps the dip represented a random score fluctuation attributable to a limited sample size, i.e., one or two targets of a particular size which are extremely inconspicuous. This idea is ruled out by the fairly large number of images involved, each of which was viewed by 30 subjects.

Perhaps, since target to background contrast was found to be an important factor in limiting Completeness scores, there had been a chance clustering of targets of certain dimensions and low contrast (conspicuity) values. But the dip appears for three major target family categories, and always in the same general range. These three categories represent 126 targets, or 59 per cent of the total number included in the Projected Display test imagery.

Perhaps the subjects were "set" to perceive targets of certain sizes because of the signature training they had received and, under rapid viewing conditions, merely failed to detect targets which were neither "moderately small" nor "moderately large." But the Experimental and Control group curves are similar and the Control group did not receive signature training. Moreover, the Control group viewed the imagery for substantially longer times than did the Experimental group. Therefore, the "flash" condition did not really apply. Finally, the dip did not appear for certain types of targets (Rail, Power, Port/Harbor).

In an effort to resolve the issue, Completeness scores were recalculated in terms of the percentage of the display area encompassed by the target area. Figure 2-14 shows the same monotonic increasing function previously noted for Port/Harbor, Rail, and Power targets. And Figure 2-15 presents results of the plots for Military, Industry, and Airfields. This latter graph evidences the same "rise-dip-rise-dip" pattern which was previously observed, and the same substantial overlap for the three curves.

It was at this stage that the analyses were extended to the Light Table Test. Fewer targets were involved (83 versus 215 for the Projected Display Test). However, it was felt that by grouping targets in terms of size ranges, sufficient stability might be obtained to permit construction of meaningful graphs. Figure 2-16 presents a comparison of the total Projected Display Test results (all target types combined) and the total Light Table Test in terms of Completeness scores achieved for targets encompassing various percentages of their respective display areas. Despite the substantial difference in display modes and despite the substantial difference in target sizes (lengths) on the two displays, the same dip occurs within the same percentage area. Perhaps at this time it should be pointed out that differences of 20 points in Completeness scores are statistically significant beyond the .01 level of confidence.

The comparison presented in Figure 2-16 made use of only the "larger" targets in the Light Table imagery (those encompassing from 1 to 100 per cent of the display area). Plots were also prepared for targets encompassing from .01 to 1.0 per cent of the display area, and these revealed a similar depression in Completeness scores. The full range of target areas is included in Figure 2-17, which presents Completeness scores achieved by the combined groups only for the Light Table Test.

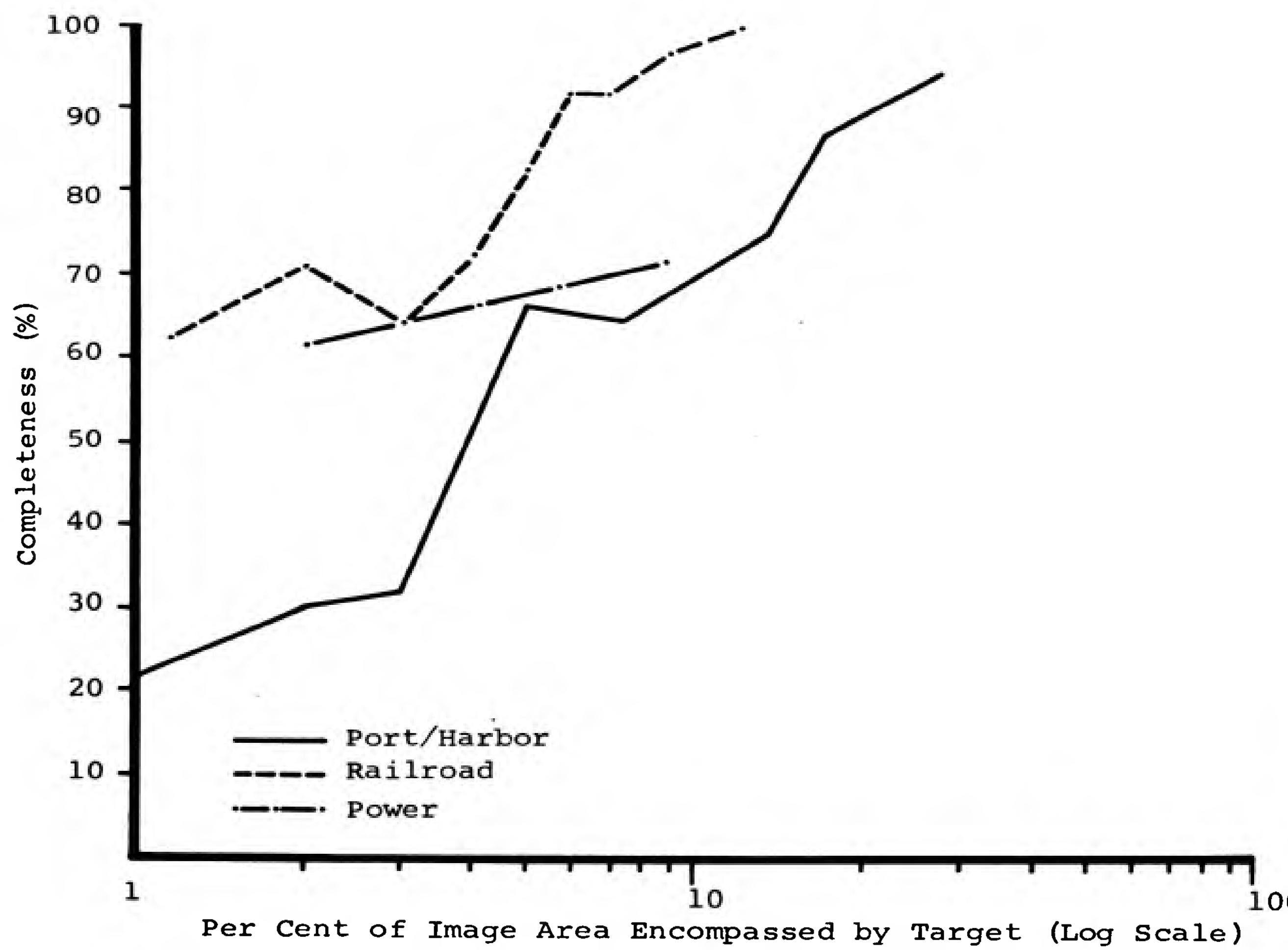


FIGURE 2-14. Port/Harbor, Rail and Power Target Completeness Scores for Targets Encompassing Various Percentages of the Displayed Area.

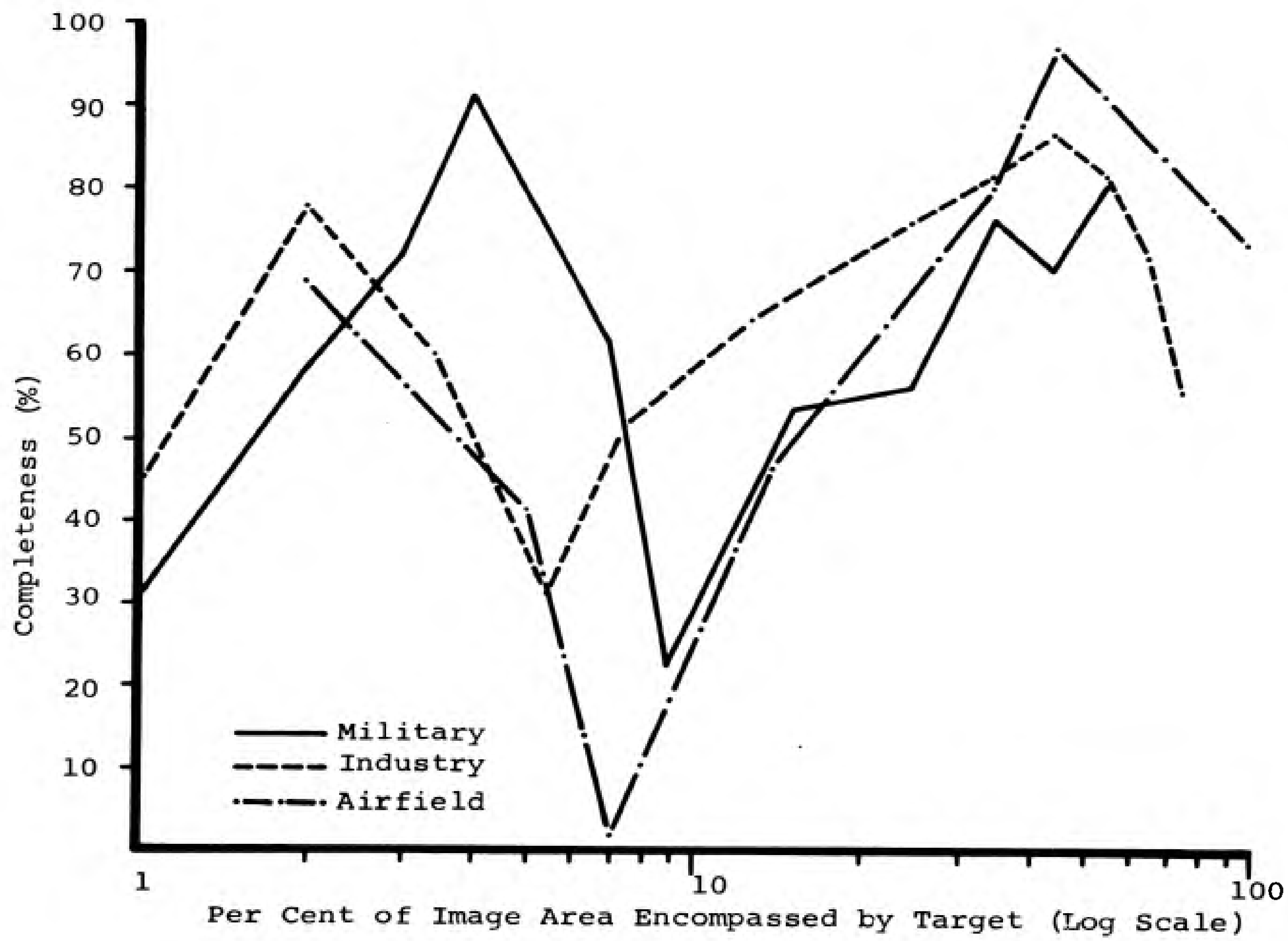


FIGURE 2-15. Military, Industry, and Airfield Target Completeness Scores for Targets Encompassing Various Percentages of the Displayed Area.

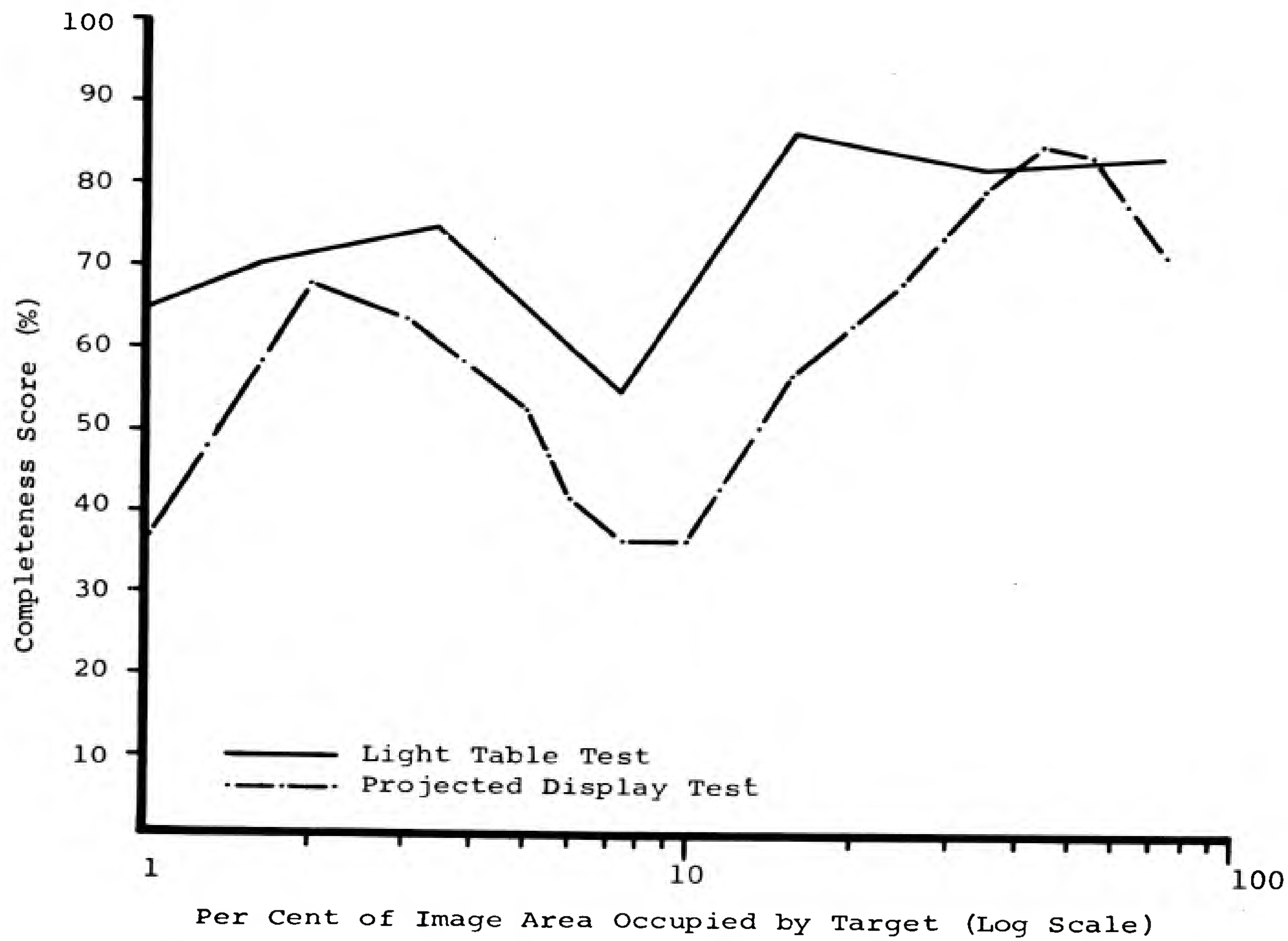


FIGURE 2-16. Comparison of Completeness Scores for Two Types of Displays in Terms of Display Area Encompassed by Various Targets.

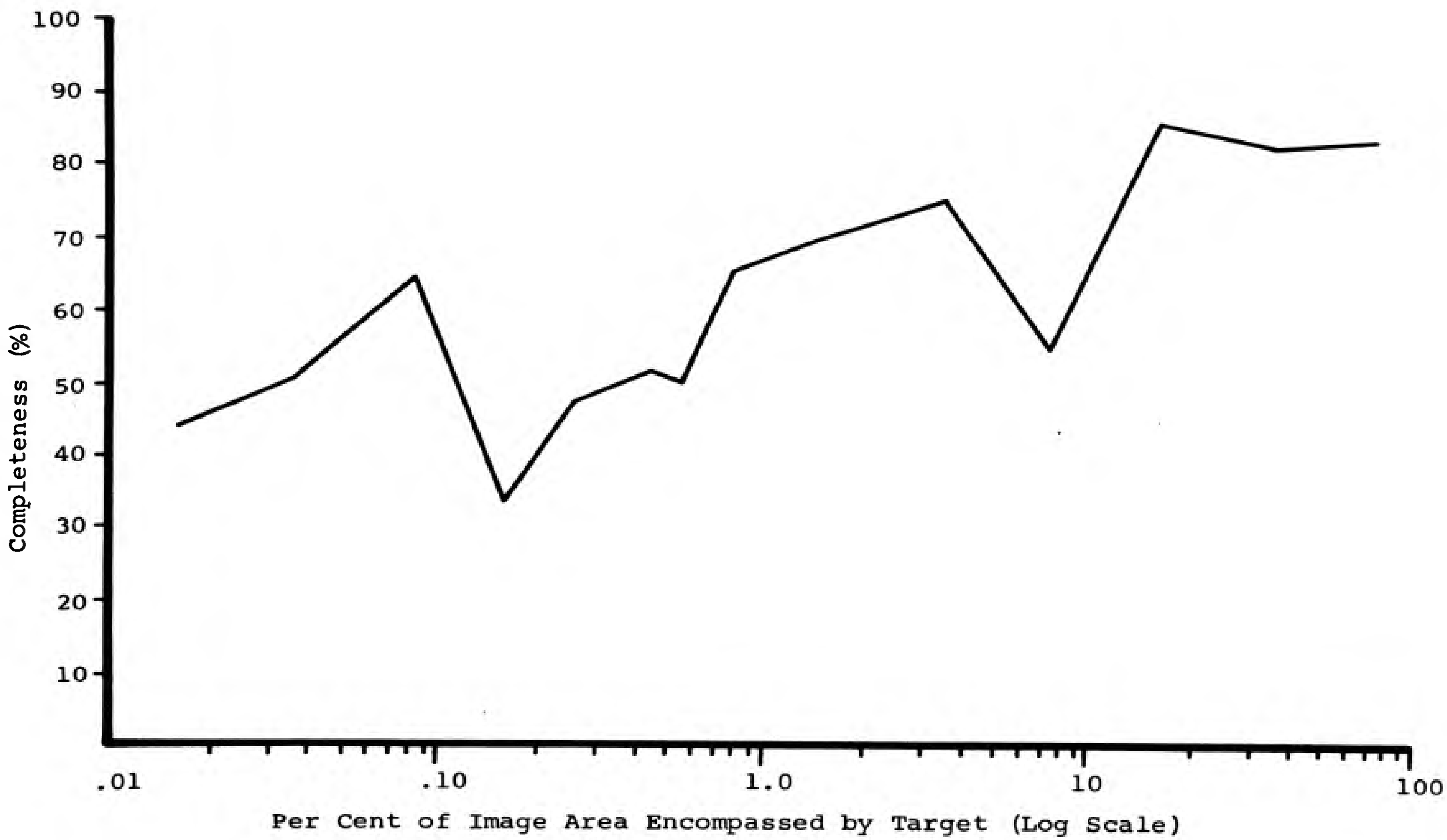


FIGURE 2-17. Completeness Scores for the Light Table Test for Targets Encompassing Various Percentages of the Displayed Area.

SECTION 3

CONCLUSIONS AND DISCUSSION

A. INTRODUCTION

This portion of the report summarizes the findings which were previously discussed and considers the implications of the findings for current operational and interpreter training practices. A section is also included which suggests certain implications of the findings for additional research efforts.

The Conclusion section does not deviate from the factual data presented in the "Results" portions of Part 1 and Part 2. However, in the Discussion section we have elected to present certain ideas for which empirical support may not be fully available. To a large extent the Discussions are based upon the study findings, but the observations of the experimental results are occasionally tempered by subjective impressions gained during conduct of the research.

The conclusions are presented below in terms of their applicability to the training program evaluation or the assessment of the effects upon performance of certain imagery and display variables.

B. CONCLUSIONS

1. The training Program.
   - The four-day training program resulted in significantly improved performance among the Experimental group of experienced interpreters compared with a control group matched in terms of experience and prior training.

- The group receiving rapid recognition training (Experimental group) was seven per cent more accurate and ten per cent more complete in their identifications than were the Control group, and they provided their responses in one-half the viewing time needed by the Control group.
- The trained (Experimental) group provided more correct responses and made fewer target omissions than did the Control group.
- There was a slight tendency for the trained (Experimental) group to report a larger percentage of targets which were not actually present in the imagery. This one to two per cent error tendency occurred with the more complex imagery.
- The Experimental group performed significantly better than the Control group in terms of Accuracy percentage for large scale imagery. The two groups performed equally for medium and small scale imagery.
- The Experimental group performed significantly better than the Control group in terms of Completeness percentage for large scale and for medium scale imagery. The two groups performed equally for small scale imagery.
- In terms of information processing capabilities, the Experimental group transmitted more information than did the Control group, and at twice the rate.
- Target signature training was at least partially responsible for the Experimental group's relatively superior performance, although not all target signatures could be presumed to function equally well in promoting the differences between the groups.

- The effects of the rapid recognition training using a rear projection display also applied when the groups were tested under more customary viewing conditions (light table display). Accuracy and Completeness scores were equivalent for both groups but the Experimental group performed the task significantly more rapidly than did the Control group (25 per cent time reduction).
- Analyses of the criterion subtests indicated a fair degree of reliability for the measures, and confirmed the independence of criterion measures of Accuracy, Completeness, Efficiency, and Time.
- Analyses of the relation between scores achieved on the Projected Display and the Light Table test indicated significant relationships for measures of Accuracy, Completeness, and Efficiency.
- Analysis of performance during the training course indicates that the four-day training course cannot be shortened. However, since there was no indication that a "plateau" had been reached, the course might profitably be lengthened. The Accuracy and Completeness scores were not so high as to rule out further improvement.
- None of the biographical data of experience or prior training served as predictors of performance on any of the criterion measures.

2. Image and Display Variables
   - Image scale is a significant variable in influencing completeness of interpretation, and interacts with target type in this regard.

- Target-to-background contrast (as measured by visual conspicuity ratings) is a significant variable in influencing completeness of interpretation, and this also interacts with target type.
- Image conplexity (as measured by number of different targets per image)is a significant variable in influencing completeness of interpretation, but the effects are reduced by the training program. The Experimental group achieved substantially higher Completeness scores at all levels of image complexity but their Accuracy was impaired at the more complex imagery levels.
- Both image scale and image complexity affected interpreter Efficiency scores but the effects were minimal for the Experimental group.
- The machine-paced Projected Display mode was far superior to the operator-paced Light Table display in terms of time required per response. Although the groups achieved comparable Accuracy and Completeness scores for each display mode, both groups performed more rapidly under the brief exposure viewing conditions established during the Projected Display Test than they did on the Light Table Test which involved continuous display of the imagery.
- Target size, either absolute length of the longest dimension or the per cent of image area encompassed by the target, interacts in an irregular manner with Completeness scores for major target types. This finding should be further investigated.

C. DISCUSSION

In the section which follows, the various study findings are considered in terms of their implications for photointerpretation operations, interpreter training, and additional research.

1. Implications for Operations. Although it is seldom that a feasibility study provides sufficient factual data upon which changes in operational procedures can reliably be based, we feel that the present study findings strongly suggest action concerning imagery display modes.

a. Display Mode. The study indicated comparable levels of Accuracy and Completeness for either Light Table or Projected Display interpretation tasks. In these respects, the displays are equal, in spite of lack of subject familiarity in using a rear projection viewing arrangement. However, with respect to time required, the briefly-exposed images on the Projected Display consistently proved to be a superior viewing mode. Consequently, it is with some confidence that we can recommend that consideration be given to the phasing over to this type of display mode. It is recognized that film processing costs or hardware costs may substantially increase over present allowances. The dollar cost must be evaluated against the five to ten times greater efficiency which is demonstrably possible through use of the projected display as it was employed in the present study.

2. Implications for Training. The present study has indicated that substantial improvements in interpreter performance can be achieved from a relatively brief course of instruction in rapid identification techniques. Although the study has demonstrated the feasibility of the approach, it was not designed to "test the limits." Consequently, a number of research issues remain, and these are discussed in the following section.

For the most part, the remaining research issues are concerned with achieving maximum benefit from the training. They do not preclude the immediate initiation of a training program patterned somewhat after the one developed during this study effort. Therefore, the present section is devoted to several suggestions concerning the design and conduct of such a training program within a military training environment.

a. Training Method.

(1) Displays: The reported research effort employed two Projected Display formats: the rear projection screen for individual testing purposes, and a wall projection for group training purposes. The size of the wall display and the subject viewing distances were approximately adjusted so that the visual angle subtended by the wall display and by the rear projection screen would be similar (36 degrees). The group size did not exceed eight men during any single training session; therefore, viewing distances for all subjects could be approximately the same.

For military training purposes, both types of projected display are recommended. The group display should be limited to fairly small groups, and has the advantage of economy of training time. The individual display possibly provides greater motivational benefits. For either type of display, reinforcement slides can be included in the image cartridges, thus automatically fulfilling certain of the basic requirements of a learning situation.

For the individual viewing configuration, equipment no more complex than that used in this study would seem to be needed* The gear has the dual advantage of economy, and flexibility of operation. The most rapid shutter speed possible is only 0.1 seconds, but exposure durations lower than two seconds are not required for any operational situation, nor are they desirable during training. We believe that the possible gains that could accrue in speeding up work tempo would be more than offset by the feelings of frustration that would be instilled in the

*See Appendix C for wiring modifications data.

students if they were required to identify targets exposed for excessively brief time periods.

(2) Curriculum: The course content of a military program will, naturally, vary somewhat from that included in the present study. Consequently, a need may exist for the generation of additional graphic target signatures. At present, the goal should be that of simplicity and the signatures should not be overly specific. The reason for this suggestion is that there are far fewer identifiable geometric forms than there are target names, and graphic "signatures" can therefore easily be developed which will be good identifiers for two utterly different target types. Such "look-alike" problems now exist and could be intensified unless care is taken in the preparation of training materials.

Graphic signatures which appeared to introduce important improvements in Experimental group scores are contained in the Appendix to this report. These, together with the description of the method used for their generation, may serve as useful guidelines to the production of additional training materials.

With respect to the speed, or scanning, training, it is suggested that training sessions be programmed over a substantially longer training period, rather than being concentrated as in the present study. Variations could also be introduced, such as training in target detection: i.e., requiring a report if a target (any target) is perceived in the briefly exposed display. This technique could later be refined so that a response is made only if certain types of targets are detected, etc.

(3) Sequencing: One point upon which most of the Experimental group subjects were in nearly unanimous agreement concerned the most appropriate time to introduce rapid recognition training. The concensus was that it should be started

early in the initial interpreter training program. We see no reason to disagree with this other than to point out that the training improved the performance of the experienced interpreters who served as subjects in this study. There would seem to be little reason to limit this type of training to photointerpretation trainees, although the problems involved in administering refresher training to operational personnel sometimes outweigh other considerations.

3. Implications for Research. The research project raised questions relevant to seven possible future study efforts. In some instances these represent application of the study technique (rapid recognition training) to related imagery formats. In other instances, additional study is necessary to clarify some of the relationships which were revealed by the present project. The seven potential projects, which are discussed in some detail below, are as follows:

- Study of the Task and Display Conditions for Which Flash Interpretation is particularly appropriate.
- Study of Display Modifications and Reporting Methods to Extend Operational Usefulness of Flash Interpretation.
- Achievement of Maximum Benefits from Rapid Recognition Training.
- Perceptual Studies Concerning Target Types.
- Extension of Training Method to Tactical Imagery.
- Extension to Include Non-Conventional Photographic Imagery.
- Extension of Graphic Signatures.

a. Study of the Task and Display Conditions for Which Flash Interpretation is Particularly Appropriate. Rapid photo-interpretation, like rapid reading, cannot be meaningfully discussed apart from a consideration of its purpose and its subject matter. Speed reading techniques have repeatedly demonstrated impressive time gains when the reader's intent is to review fairly standard, highly redundant, narrative materials such as novels or reports. The techniques do not, however, apply if the task is, say, to study a calculus text. Rapid scanning, per se, can be expected to effect little or no reduction in overall time required when the bulk of the task is devoted to the understanding of obscure relationships among highly abstract symbols. In fact, accuracy, (i.e., understanding) may be impaired if the act of rapid scanning results in misinterpretation or misidentification of the symbols.

Similarly, rapid scanning of pictorial imagery can be expected to have relevant applications for certain tasks and may be entirely irrelevant for others.

If the task is that of preliminary screening of imagery, it might be expected, on the basis of related research, that substantial time benefits would accrue. This has, in fact, been demonstrated in the present study. The benefits might not necessarily be expected to hold if the task were to analyze a target area in terms of its function or its level of activity or its threat potential. However, the present study indicated that even for a preliminary screening task, rapid scanning may not be an unmixed blessing. If the target categories of interest are small and/or camouflaged they may be missed. Or if decoy targets have been constructed, they may tend to mislead an interpreter who has not taken the time to detect the subtle cues which may reveal the true nature of the perceived object. The likelihood of either of these interpretation errors occurring may be expected to increase as image quality is degraded. These suppositions are supported by the study results which indicate a higher percentage of erroneous identifications occurring in the Experimental group. The seriousness of this type of error must be weighed against the gains in terms of interpretation time and efficiency.

One of the major benefits to be derived from a study of speeding human processes is not to establish feasibility, but to define the conditions under which speeded performance yields useful gains. It is likely that in accomplishing certain interpretation tasks, or in viewing certain kinds of images, rapid scanning is contraindicated. When faced with these situations, the interpreter may elect an alternative viewing strategy. If the parameters are pre-established, the interpreter may select the optimum viewing strategy with less likelihood of error.

Some illustrative hypothetical cases may clarify this point.

If it has been determined that for rapid detection of certain target types a scale value of 1:5,000 or larger is required to achieve a 90 per cent detection probability, and that detection probability falls to, say, 50 per cent at scale values of 1:10,000, an interpreter would have objective data upon which to base his selection of viewing strategy. If, for example, his imagery is 1:7,000 he may judge his (interpolated) probability of detecting the desired targets to be sufficiently high to permit rapid scan in view of a high requirement for speed and a relatively low penalty (in this hypothetical instance) for error. If, on the other hand, the error penalty is high, the time requirements are urgent, and the imagery scale is very small, rapid scanning would not be suitable and the desirability of assigning additional interpreters might be indicated.

If the above assumptions are correct, the remaining task is to determine the relevance of specified variables to completeness and accuracy criterion measures under conditions of speeded performance. A major problem in such a task concerns not only establishing the influence of each of a large number of variables but also determining their interactive effects. As the number of variables increases, a corresponding increase in appropriate test materials and test subjects is required. And both appropriate imagery and qualified subjects are rare. The individual and collective influences upon PI performance of a substantial number of variables needs to be determined: type of target, type of sensor, scale, resolution, contrast, density, acutance, etc. The present study has demonstrated the relevance of certain of these variables, but this can be regarded only as an initial step toward the goal of understanding the perceptual problems involved in accomplishing the image interpretation function.

The research need (and this extends beyond the simple determination that image quality affects interpreter performance) is to quantify the relationships to the extent that probabalistic estimates which are based upon certain quality measures can be made of interpretation completeness for specific target types.

b. Study of Display Modifications and Reporting Methods to Extend Operational Usefulness of Flash Interpretation. Two of the major variables involved in the task of the photointerpreter in discriminating significant shape patterns (targets) are speed and accuracy. An elemental issue, basic to advancing interpreter capability in accomplishing his task, involves an appreciation of the ways in which speed and accuracy co-vary.

They may co-vary, first, as a function of variations of the relative importance of errors versus time. Generally, there is a substantial negative correlation which indicates that accuracy is purchased at the expense of time. In accomplishing most motor tasks for example, speeded performance degrades accuracy and lengthened performance time improves accuracy. In the present study, of 30 intercorrelations between various time and accuracy measures, 27 were negative in sign. However, there is some evidence to indicate that, for certain perceptual tasks, errors may increase with increased time available. If accuracy is defined solely in terms of the number of correct responses per unit time, then the inverse relation between speed and accuracy for motor tasks may continue to hold for perceptual tasks. If, however, accuracy is expressed as a ratio of correct responses to total responses made, the nature of the speed-accuracy relationship for perceptual tasks is probably curvilinear.

A second way in which speed and accuracy may co-vary is as a function of individual differences. Generally a low positive correlation is found among abilities. Individuals who are most accurate in accomplishing a task also tend to perform the task somewhat more quickly than those whose accuracy scores are low.

Thirdly, speed and accuracy may co-vary as a function of display design. A demonstrable relationship exists between the methods used to display information to a subject and the suitability of his task performance. Changes in the design of a display which improve speed of motor performance will generally also improve accuracy. There is reason to suspect (based upon present study findings) that a similar relationship would hold for display mode and speed of perceptual performance.

In considering display variables which may be improved, a distinction is required between aspects of visibility and legibility. Visibility refers to a target's capacity for being seen and legibility refers to a target's capacity for being identified. In a photointerpretation context these translate to the functions of detection and identification. (Only the identification aspect was tested in the present study and is expressed as a "Completeness" score; i.e., completeness of identification). Although a low positive correlation generally obtains between measures of accomplishing the functions of detection and identification, typically no relation is found between measures of visibility and criteria of speed or accuracy. The gains to be expected in the criterion measures are largely a function of improvements in legibility, and reference to the curves for Completeness as a function of target-to-background contrast indicates that the potential for improved performance is considerable.

Legibility (identification) improvements may be effected by improving the definition of margins or borders of a target or by the use of color. Other changes which have been found to improve legibility (of alphanumeric symbols) involve line width and spacing, height/width ratio and line inclination. Some of these may be translatable into methods for improving the "legibility" of pictorial displays.

One further method to gain speed in task performance involves reducing the requirements for precision. Time can be gained by reducing the information on the display (hence, the number of discriminations required) or by reducing the information which the subject must transmit.

Display simplification can occur from an elimination of irrelevant signals. This may be effected on projected displays by various forms of level slicing or image enhancement.

Reporting simplification can occur by requiring less specificity in target identification. This may be accomplished by developing a hierarchical or familiar classification of objects of interest similar to that employed in the present study. If the subject is required to report only at a very gross target family level (say, among industry, shipping, airfields, and missile sites) the amount of information transmitted per response is low but his response time will also be rapid. The degree of improvement may be evaluated in terms of a contrasting situation which requires the observer to identify specific subclasses of the target families, e.g., a certain type of missile site, a specific type of industry, etc. In this latter instance, the amount of information transmitted is higher but an increased amount of time is required in order to make the necessary additional discriminations.

Some exploratory work has already been accomplished by various groups to determine whether electronic processing of pictorial displays will actually improve interpreter performance, and fairly elaborate consoles are currently being built to extend these test efforts. Because of the relationships revealed by the present study, it is recommended that future studies include consideration of target type and target size in addition to the several quality variables of scale, density, contrast, acutance, resolution, etc. It may be that the value of electronic image processing is specific to certain target types and either unnecessary or of no help for others. In addition, it is suggested that the criterion measures of interpreter performance not be limited to elemental detection levels, but provide for the measurement of "depth of completeness" of interpretation.

c. Achievement of Maximum Benefits from Rapid Recognition Training. As has been noted, the goal of the present research effort was to develop and test a program of training in techniques for the rapid, accurate, and complete extraction of intelligence information from pictorial imagery. For various practical reasons, the program which was developed was administered in four consecutive working days with a fifth day reserved for evaluation purposes.

During the four-day training period the Experimental group subjects were exposed to a concentrated schedule of instruction, drill, and periodic testing. The results of tests administered during and immediately after the training program indicated an increasing capability level for all criterion measures. These results, it will be recalled, were obtained for experienced interpreters, not for photointerpretation students.

Several questions occur. For example, the argument concerning the relative merits of spaced practice and massed practice has a long and honorable history in psychology. Would the improvements which were obtained through the concentrated course of instruction of the present study be increased if the training program were extended over a longer period of time but with no increase in content?

Since the Experimental group was apparently still improving at the time the final test was administered, does this test reveal the peak performance which can be expected? Or is there reason to believe that additional improvements could be achieved if more training and practice were to be provided? In short, the present study established the feasibility -- it did not determine the limits -- of training in rapid interpretation.

A related question concerns the time duration over which the improvements can be expected to remain. It was shown that the display mode had an effect upon interpreter performance. Since the Experimental group subjects will return to their more customary (light table) displays, will the training course effectiveness tend to dissipate over time? The study findings suggest a qualified "yes." The trained group performed far less efficiently on the Light Table test than they had on the Projected Display test (but this may have been due to the machine-regulated display duration on the Projected Display test). It might be surmised that the effects of the "speed" portion of the training program would attenuate when the subjects revert to a customary viewing mode. However, the Experimental group, despite a loss in performance efficiency on the Light Table test, still was significantly superior to the Control group. It might be surmised that those portions of the program concerned with speeding a personal work tempo could be expected to have only short term improvement effects (unless periodic reinforcement is provided). Those portions of the training program dealing with target signatures may be expected to have fairly long term effects. These are suppositions, however, and might well be subjected to test.

d. Perceptual Studies Concerning Target Types. Although the phenomenon of the decrease in Completeness scores as target size increases through mid-range values is not believed to be of immediate operational significance, it is recommended that it be subjected to additional test. It seems likely that the result is an artifact, yet all "rational" steps, from inspection of slides to further analytic comparisons fail to reveal why the phenomenon should appear for such a large percentage of the targets, and for only certain target types. It is possible that the effect is introduced by the scoring formula, i.e., the particular weighting system employed, in conjunction with such variables as image scale and image "complexity." A re-analysis of the data, scoring only at the highest level of abstraction, would serve to clarify this situation.

If a replication study should confirm the finding, then immediate efforts would appear to be indicated to design and evaluate a display console which would permit variable magnification of selected segments of a display. If the tendency does exist, such a viewing console would seem to be one promising means of correcting the condition.

e. Extension of Training Method to Tactical Imagery. During the initial imagery screening effort, attempts were made to acquire a quantity of tactical imagery, particularly imagery containing armor, gun emplacements, bivouac areas, convoys, weapons carriers, supply depots, and troop concentrations. Some imagery was obtained, but too little to include in the training and testing programs.

It would seem that the need for rapid appraisal of imagery for possible tactical targets is at least as great as the need to screen for strategic targets. It would also seem that training results established for large strategic targets might not necessarily be directly extrapolated to tactical imagery. (For example, except for vehicle identification, the concept of target "signatures" might be rather difficult to convey.)

Since the present study has established the feasibility of the approach, a logical extension would appear to be toward determining the training relevance for tactical imagery. Such a study should be oriented not only towards photographic transparencies but also toward near real-time interpretation.

f. Extension to Include Non-Conventional Photographic Imagery. The viewing of non-conventional photographic imagery (infrared and side-looking radar) would appear to occupy about the same position as the tactical imagery mentioned above. Although it is unlikely that either of these types of sensor imagery would be rapidly screened, the present study has indicated an area for further research which is most relevant for these forms of imagery. Specifically, both infrared and radar imagery may be viewed in transparency format and, in this format, they are customarily viewed on a light table. It is interesting to speculate whether a briefly exposed projected display would have the same effect in improving interpreter efficiency in viewing non-conventional photographic sensor imagery as it had for the imagery employed in the present study.

In addition to the display mode issue, relevant areas of research for these sensors would appear to include the feasibility of developing graphic signatures. Because of the vastly different characteristics of the sensed energy it is likely that the types of signatures would be different from the simple graphic forms employed for the conventional photographic imagery. Development of the signatures would necessarily require a carefully conducted testing effort to assess their merits.

g. Extension of Graphic Signatures. Finally, the present study has provided information concerning the merits of graphic signatures in promoting rapid identification of certain types of targets. All targets did not evidence the same improvement in Experimental group performance. This finding has several implications. On the negative side, it might prove to be of value to initiate a graphic signature development study, to investigate alternative graphic "aids-to-recognition" for targets on which the Experimental group did not achieve significantly.

On the positive side, since the merit of certain signatures was established, it might be helpful to consider their application to interpretation functions other than target detection. For example, although the recognition or identification process has been speeded, the reporting process has not. A mechanical keying system, coded in terms of graphic signatures might have considerable merit as a rapid reporting device for relaying information concerning identified targets. Similarly, the graphic signatures might have a useful application in the structuring of photointerpretation keys, and other imagery repositories.

REFERENCES

Attneave, F. "Some Informational Aspects of Visual Perception," Psychol. Rev. 1954-61, 183-193.

Attneave, F. Applications of Information Theory to Psychology, Henry Holt and Company, N. Y., 1959.

Alluisi, E. A. "Conditions Affecting the Amount of Information in Absolute Judgements," Psychol. Rev. 1957-64, 64-97.

Bice, R. C. "An Exploratory Image Interpretation Experiment," Part I, June 1958 and Part II, November 1958, Army Ballistic Missile Agency - Cont. DA-36-034-ORD-2677.

Enoch, J. M. "Natural Tendencies in Visual Search of a Complex Display," Visual Search Techniques, NAS-NRC, Public. 712, 1960.

Erikson, C. W. "Location of Objects in a Visual Display as a Function of the Number of Dimensions on Which the Objects Differ," J. Exper. Psychol. 1952, 44, 56-60.

Luborsky, L. "Aircraft Recognition: I. The Relative Efficiency of Teaching Procedures." J. Appl. Psychol. 1945, 29, 385-398.

Luborsky, L. "Aircraft Recognition: II. A Study of Prognostic Tests." J. Appl. Psychol. 1945, 29, 449-457.

Minturn, A. L. & Reese, T. W. "The Effect of Differential Reinforcement on the Discrimination of Visual Number." J. Psychol. 1951, 31, 201-231.

Verplanck, W. S. & Blough, D. S. "Randomized Stimuli and The Non-Dependence of Successive Responses at the Visual Threshold." J. General Psychol. 1958, 59, 263-272.

NRC Committee on Vision. Form Discrimination as Related to Military Problems, 1957.

NRC Committee on Vision. Visual Search Techniques, 1960.

NRC Committee on Vision. Vision Research Reports, 1960.

NRC Committee on Vision. Visual Problems of the Armed Forces, 1961.

APPENDIXES

	SUBJECT RESPONSE		
KEYED ANSWER	3 POINTS FOR	2 POINTS FOR	1 POINT FOR
MS		MS	MIL MB
MB		MB	MIL MS
ECS	ECS	AMMO EXP	MIL CHEM
AMMO	AMMO	ECS EXP	MIL
EXP		EXP AMMO ECS	MIL
IRBM	IRMB	NIKE ICBM MISS	MIL
ICBM	ICBM	NIKE IRBM MISS	MIL
ELEC		ELEC RADAR RS ANT	MIL
NIKE	NIKE	ICBM IRBM MISS	MIL
ANT	ANT	RS RADAR ELEC	MIL
MISS		MISS ICBM NIKE IRBM	MIL

SUBJECT RESPONSE

KEYED ANSWER	3 POINTS FOR	2 POINTS FOR	1 POINT FOR
RS	RS	ANT ELEC	MIL
RADAR	RADAR	ELEC ANT	MIL
MIL			MIL EXP ELEC RADAR RS ECS ANT MB MISS MS ICBM IRBM AMMO NIKE
POL		POL	REF CHEM
HF	HF	LF FAB	PROC STEEL IND
LF	LF	FAB HF	IND
FAB		FAB LF HF	IND STEEL
REF	REF	CHEM POL PROC	IND
CHEM	CHEM	REF POL PROC	IND
STEEL	STEEL	PROC	IND FAB HF

	SUBJECT RESPONSE		
KEYED ANSWER	3 POINTS FOR	2 POINTS FOR	1 POINT FOR
PROC		PROC STEEL CHEM REF	HF FAB IND
IND			IND FAB PROC HF STEEL LF CHEM REF
TY		TY	MY RR
RH		RH	RR
MY		MY	RR TY
RR			RR MY TY
PM		PM	CAF LF AF MAF
PAF		PAF	MAF CAF AF
CAF		CAF	PAF MAF AF PM
MAF		MAF	CAF AF PAF PM

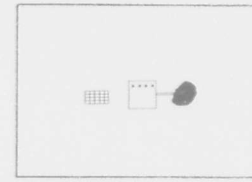
	SUBJECT RESPONSE		
KEYED ANSWER	3 POINTS FOR	2 POINTS FOR	1 POINT FOR
AF			AF MAF PAF CAF PM
MSY	MSY	SY CSY	MP/H CP/H P/H
CSY	CSY	SY MSY	MP/H CP/H P/H
SY		SY CSY MSY	MP/H CP/H P/H
MP/H		MP/H	CP/H P/H
CP/H		CP/H	MP/H P/H
P/H			P/H MP/H CP/H
APP		APP	PWR TEP
HEP		HEP	PWR
TEP		TEP	PWR
PWR			HEP TEP APP PWR

Thermal Electric Power

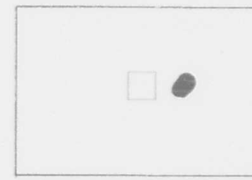
Signature Elements: Coal pile; power station; stacks; conveyor; transformer yard

Speed Signature Elements: Coal pile; power station

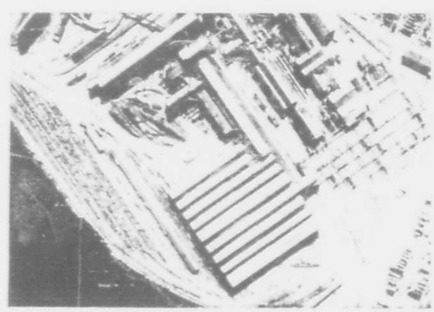
Associational Features: Nearness to water; railroad



Target Signature



Speed Signature

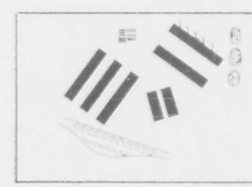


Steel Plant

Signature Elements: Long, low rectangular buildings; stacks; open storage area; ore piles; railroad and/or marshalling yard

Speed Signature Elements: Long, low rectangular buildings; railroad/marshalling yard

Associational Features: Nearness to water; located in non-residential area; lack of vegetation



Target Signature

Speed Signature

Roundhouse

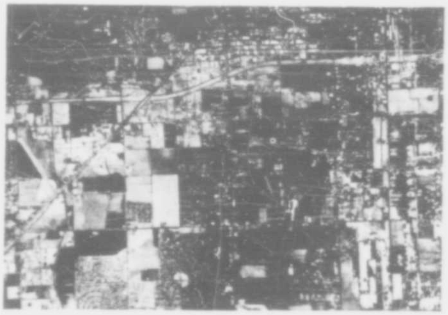
Target Signature

Speed Signature

Signature Elements: Sickle-shape; turntable; rail lines leading to turntable

Speed Signature Elements: Sickle-shape

Associational Features: Located within or adjacent to a marshalling yard



Railroad

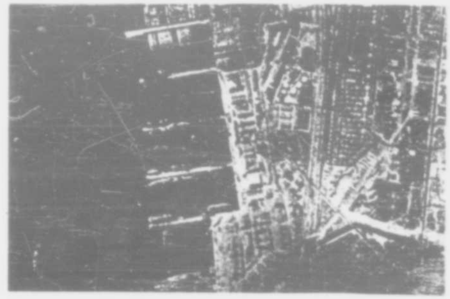
Target Signature

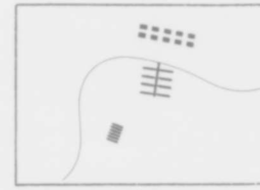
Speed Signature

Signature Elements: Long, thin line; no abrupt curves; dirt tone to track bed

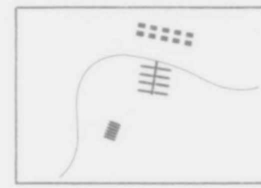
Speed Signature Elements: Long, thin line; no abrupt curves

Associational Features: If present, rail cars appear as chain-whereas autos on highway appear as dots





Target Signature



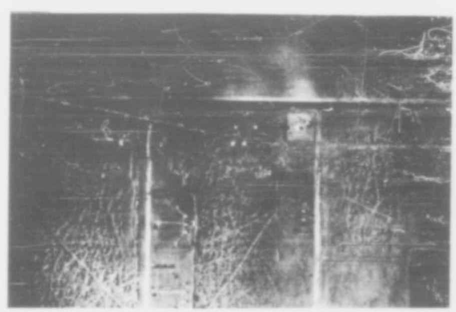
Speed Signature

Military Port/Harbor

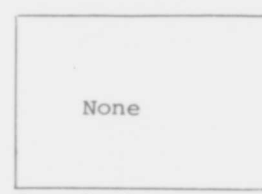
Signature Elements: Ships docked side-by-side (parallel to shore) or "nested" in mid-stream; many similar type ships; military base on shore

Speed Signature Elements: Regular arrangement of ships and buildings

Associational Features: None



Target Signature



Speed Signature

IRBM

Signature Elements: T-shaped guidance unit; arrow shape of missile on launcher

Speed Signature Elements: None

Associational Features: Remote location; secure area

Military Base

Target Signature

Speed Signature

Signature Elements: Regular arrangement of barracks and buildings; parade ground

Speed Signature Elements: Regular building arrangement

Associational Features: Remote location; secure area (fence, perimeter road); possible access to rail facilities

Military Supply

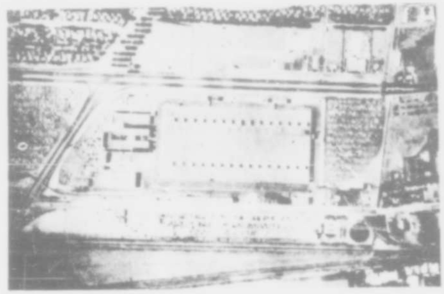
Target Signature

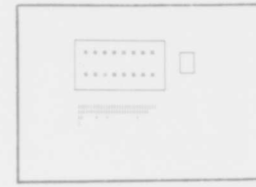
Speed Signature

Signature Elements: Many large, rectangular buildings regularly arranged; striped roof effect

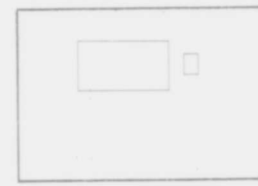
Speed Signature Elements: same as above

Associational Features: Access to rail facilities





Target Signature



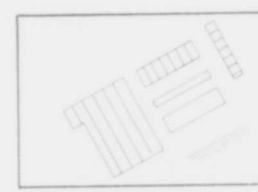
Speed Signature

Light Fabrication

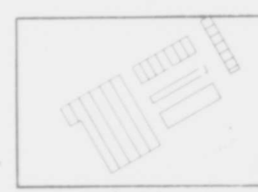
Signature Elements: Large, low rectangular building; flat roof with ventilators; smaller main office building; extensive parking area

Speed Signature Elements: Large, low rectangular building; smaller office building

Associational Features: May be located near to a residential area



Target Signature



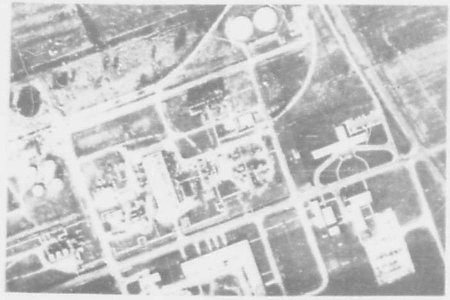
Speed Signature

Heavy Fabrication

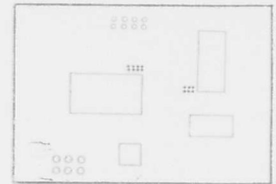
Signature Elements: Numerous buildings occupying large area; monitor roofs yielding striped effect; large parking area

Speed Signature Elements: Numerous buildings with striped roof pattern

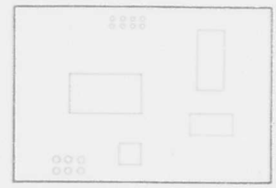
Associational Features: Located in non-residential area; access to rail facilities



Chemical Plant



Target Signature

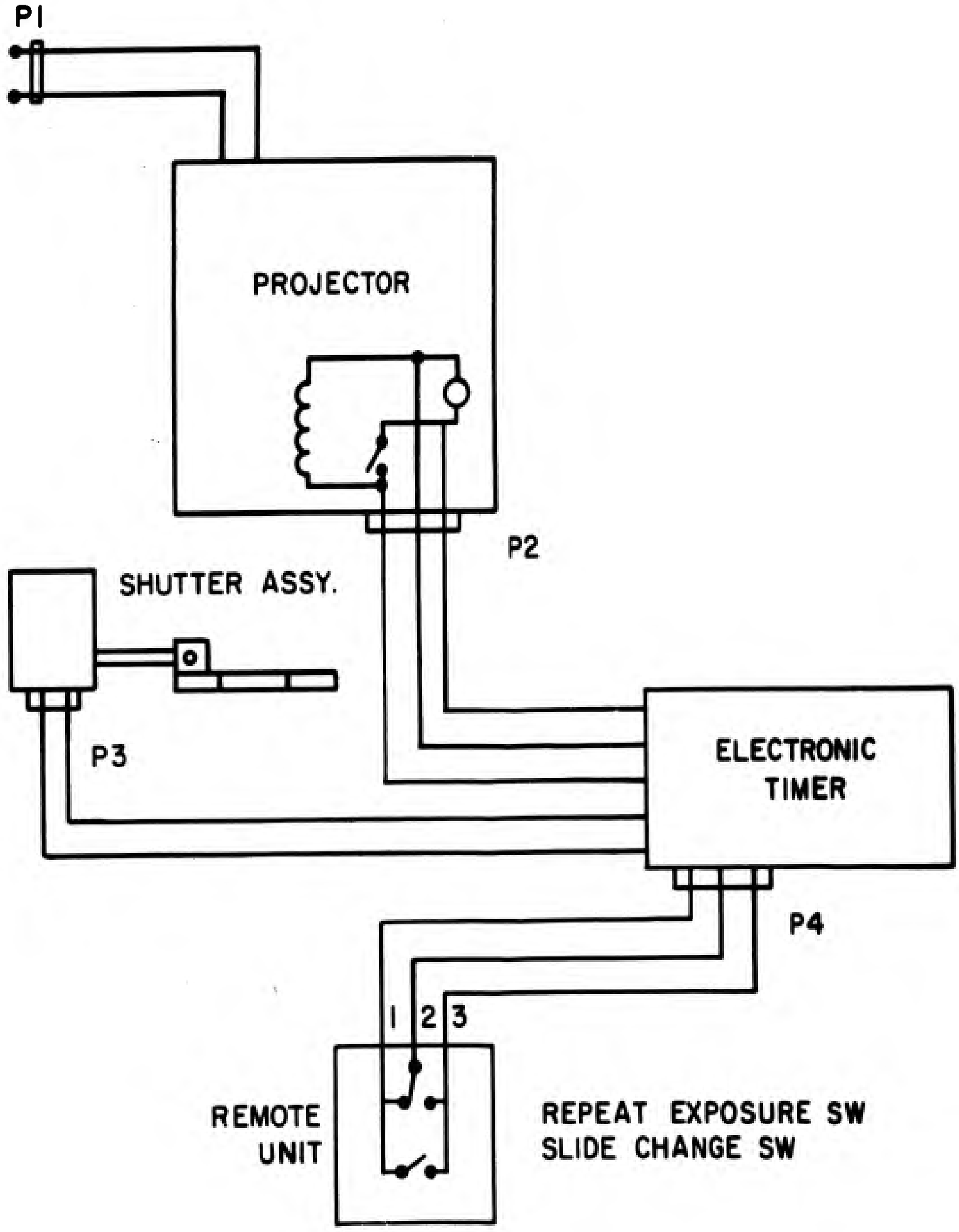


Speed Signature

Signature Elements: Variety of building shapes; POL; storage tanks near buildings; plant spread out over large area

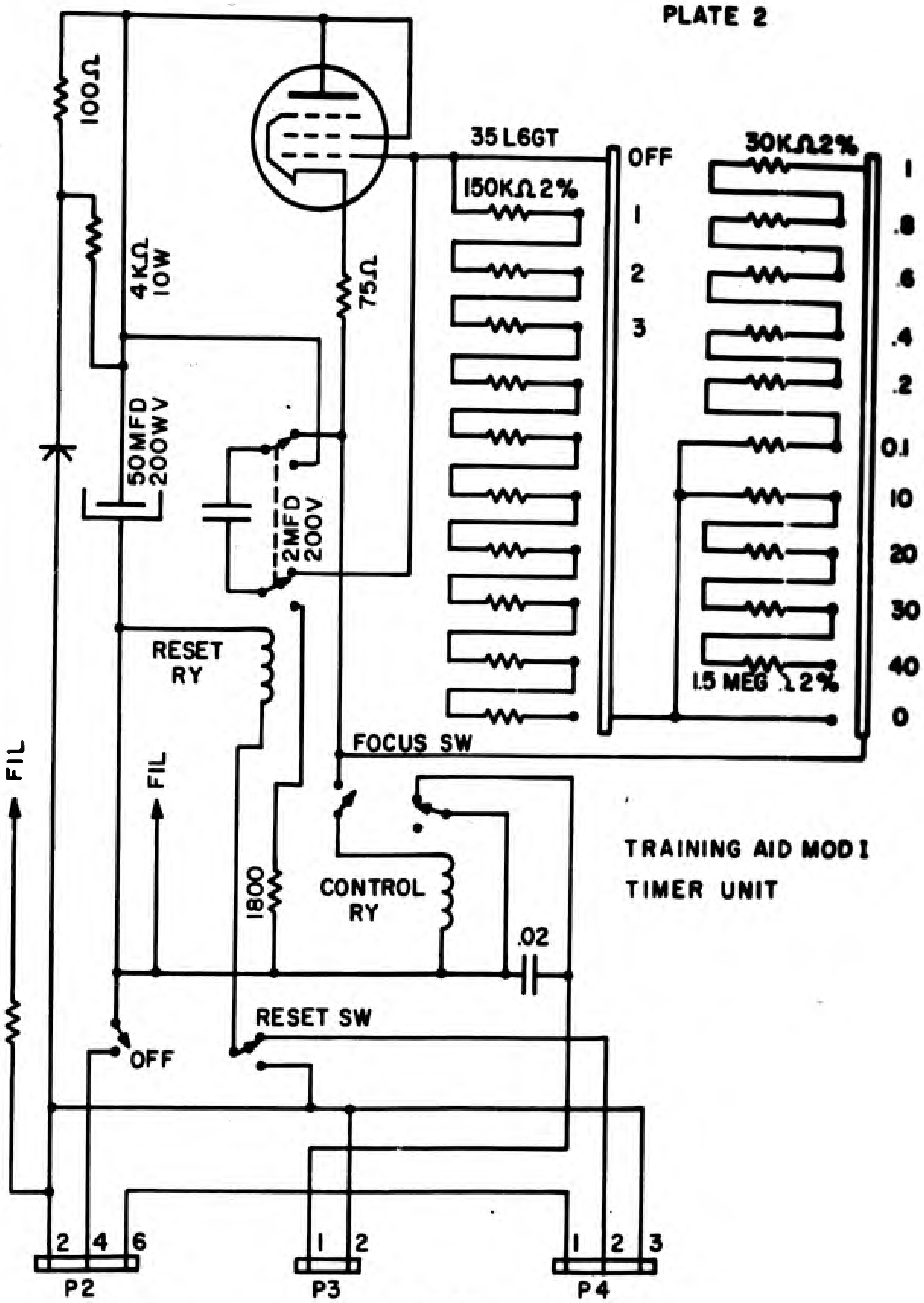
Speed Signature Elements: Variety of buildings; POL

Associational Features: Remote location; access to rail facilities



TRAINING AID MOD I
COMPLETE ASSY.

PLATE 2



TRAINING AID MOD I
TIMER UNIT

UNCLASSIFIED

END

UNCLASSIFIED