PREDICTING THE REPORTING ABILITY OF TEST SUBJECTS

By

Herbert I. Abelson

Thesis submitted to the Faculty of the Graduate School of the University of Maryland in partial fulfillment of the requirements for the degree of Doctor of Philosophy

1952

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

ACKNOWLEDGEMENTS

The development and validation of a reporting ability test was carried on as part of Contract DA 44-109qm-129 between the University of Maryland and the Office of the Quartermaster General, Department of the Army. Many individuals contributed to the successful completion of this research, and it is at the risk of oversight that the author wishes to express his indebtedness and appreciation to those whose help was especially valuable.

Dr. John M. McGinnis, Chief of the Human Resources Unit, and Mr. R. Palmer Benedict, Office of the Test Director, both at the Office of the Quartermaster General, integrated the plan of research with the specific needs of the Quartermaster Corps, and expedited the collection of data for the validation study.

Major George W. Baccus, Chief of the Survey Division at the Quartermaster Board, Fort Lee, Virginia, supervised the organization and conduct of the data collecting at Fort Lee with commendable efficiency and care. The three specialists who administered the tests, Cpl. Robert S. Andrews, Pfc. Stephen M. Sickle, and Pfc. Douglas A. Trask, manifested an insight into the research much beyond what was required of them. Mr. Elie Weeks, Chief of the Field Observer Branch of the Survey Division, Quartermaster Board, supervised the integration of the summerweight test questionnaire data to make it more meaningful.

Dr. Alan M. Kershner and James Parker supplied a number of excellent suggestions during the planning and development phases of the investigation. Reuben Shevitz analyzed portions of the data and advised on the analysis procedures.

Howard Hembree and Thomas Hussman were on hand for every crisis, and administered criticism, help, and personal therapy in equal parts.

Mrs. Janice Fish transcribed verbal reports from both the recall pretest and the validation study.

Mrs. Mary Hembree transcribed a number of verbal reports, typed the manuscript, and used every guile to maintain high motivation in the author.

The members of the dissertation committee, Dr. Arthur W. Ayers, Dr. Charles N. Cofer, and Dr. Denzel D. Smith critically examined each phase of the project; their recommendations contributed greatly to the value of the study.

The day to day guidance of Dr. Ray C. Hackman, chairman of the dissertation committee, and the author's academic advisor for the past four years, was invaluable to the conduct of this research. Any attempt at an adequate statement of appreciation at this time would be presumptuous.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

THE PROBLEM

Historical Background

The research to be reported in this dissertation is allied to research in the psychology of testimony. Many investigations in the area of testimony have been concerned with the same kinds of problems as those treated in the following chapters. It would be appropriate at this time to review some of the literature on testimony so that a better evaluation may be made of the contributions of this investigation in the light of what has already been reported in studies in testimony and related fields.

The psychology of testimony seems to have received more emphasis during the early part of this century than at any other time. Whipple (20) was concerned with situations in which the observer under controlled conditions makes a report of an event he has just witnessed, or a picture he has just seen. Whipple states that "....these tests have been found to possess a peculiar value, particularly in the study of individual differences in mental constitution and mental efficiency". Whipple further defends research in the area of testimony by generalizing that, because language is so important in our activities, any studies which cast experience in linquistic form are well worth while on that account alone.

Allport and Postman (1) comment on this period by suggesting that

> At the time Whipple wrote, psychologists found themselves attracted to the study of testimony, one suspects, because it was about the only applied field involving higher mental processes where they felt they could put their science to practical uses.

Binet (4), in 1900, was the first to propose that the creation of a practical science of testimony would be extremely valuable. Binet and Stern (18) in the decade that followed, set the pattern for investigations in this field, and were probably the two most prolific contributors to it. The problems most often emphasized in succeeding years were these (21):

1. The nature of the stimulus situation.
2. Length of exposure of the observer to the stimulus situation.
3. The elapsed time interval between exposure to the stimulus and report.
4. The form of eliciting the report.
5. Categorizing the material in the report.
6. The pattern of distortion between the report and the objective stimulus.
7. Relation of reporting ability to age, sex, and intelligence.
8. Generalizing from the analyses of the reports to the characteristics of a good courtroom witness.

The apparatus for one test (20) consisted of a cigarette, a lithograph, and a cancelled two cent postage stamp.

The observer was directed to write a description of each item in turn, and was allowed up to ten minutes for each stimulus. In another kind of test, the observer became a witness to a short scene, acted out before him, either with or without his knowledge that this was a prearranged situation. Gradually, however, the most widely used form of test was a picture or series of pictures. The claimed advantages of using pictures as stimuli were (21): they allow an exact and detailed comparison with the report; they are inexpensive, easily administered, and readily repeated; they offer experiences of varied types and complexities, and provide a situation not wholly unlike everyday life. Although the use of pictures has been criticised as being too artificial and unlike the kinds of situations dealt with by legal testimony, Whipple (21) states in this connection that "....experiment has shown that inferences drawn from the picture test have the general validity that is claimed for them".

Another variable in the investigation of testimony was the interval which elapsed between the exposure to the stimulus and the elicitation of the report. At least one investigator (2) experimented with time lapses of several days, a half an hour, and with those of a few minutes. Earlier writers (14, 5) distinguished between the kinds of mental processes involved in short and long time intervals. Meumann (14) wrote that a report elicited immediately after the material was presented was a test of "observation and its effect

as manifested in the immediate retention of complex impressions". When a considerable time has elapsed, we are dealing with "true memory". Whipple (20) classifies as description the observer's account of the experience at the time of his observation; if this account is given at any time subsequent to the observation, it constitutes report.

The manner of eliciting the report was accorded much attention, since it was believed that the way in which the observer was interrogated would determine the kind of information available for later analyses (4, 5, 14, 18, 20, 21). The two types of reports frequently used were the free description of the event and the interrogatory, or cross examination. The latter was introduced because of the opportunity to exploit suggestibility, a prime concern of these early investigators. Binet (4) listed five types of questions which could be used on the continuum of suggestibility. In the order of least to most suggestibility in the question, these were determinative, completely disjunctive (a choice between two alternatives), incompletely disjunctive (a choice between two alternatives with additional alternatives plausible), expectative ("Was this not true that...."), and implicative (asking, for example, the color of an animal, when there was no animal in the picture). Binet also included the category of the consecutive question, which was an improvised item, used to augment suggestion developed by previous questions. The number of questions in the cross examination ranged from fifteen to one hundred, with about fifty the modal number.

Analytic methods for processing the material obtained from the reports was a problem which received a great deal of emphasis. The numerous speculative attempts at its resolution are one indication of the uncertainty with which it was attacked. Binet (5, 20), for example, listed four categories into which, with varying degrees of intensity, all reports could be classified. He did not clearly indicate which types are indicative of a good reporter:

1. Descriptive - a mere catalog of the features of items.

2. Observational - more emphasis given to interrelations of the objects seen. Tendencies to interpret and conjecture.

3. Emotional or poetic - less accurate in observation, with introduction of sentiment or imaginative interpretation.

4. Erudite - uses past experience to expand on what is known about the object. Interjects bits of personal information about the object.

Whipple (21) states that the accuracy of a report depends upon both quantity and quality. As a quantitative measure, he suggests counting the number of items mentioned, or the number of questions answered. This measure constitutes the range of report. The qualitative measure is the fidelity of the statements made, designated as the accuracy of report. For any given reporter, range and accuracy are said to be inversely related. The interrogatory form of eliciting the report tends to increase range and decrease accuracy. Whipple further notes that persons, their acts, objects, and spatial relations are reported with high accuracy, while

quantities and colors are reported with low accuracy. This phenomenon he termed the process of selection.

Other psychologists were also intrigued with the pattern of distortion which could be noted by comparing the report with the stimulus picture. Allport and Postman (1) note Stern's comments on this topic. Stern said that the first distortions and omissions seem to occur during the original perception of the picture or episode, and that the event to be reported must stand alone in the subject's mind for accurate reporting to obtain. Whipple (20) says that the "description of an object is inadequate, because it is almost invariably simplified". He also states that (21), "the observer's account [because of the limitations of language], can never do justice to the continuity, complexity, and delicacy of shading of his mental experience....." And again, "....the observer not only observes, but .... he also reports..... [It is] practically certain that the report is only a partial, and often a misleading statement of the real experience".

With respect to certain of the characteristics of the reporter (21), men have been found to be 20% to 33% more accurate than women, children are in every way inferior to adults, and nothing conclusive may be said about the relation of reporting ability to intelligence.

In a later study, using over 500 school children as subjects, McGeoch (13) investigated the effects of age and sex on reporting with intelligence held constant. The subjects,

nine to fourteen years of age, described and later were cross-examined on an object card, picture, and event. McGeoch found marked differences in ability to report at the extremes of the age range, small differences between any two consecutive years. Sex differences were found to be partly a function of the material, with girls consistently giving better reports. Girls did best in describing the event, less well in describing the picture, least well in describing the object card.

A more recent use of testimony is reported by Allport and Postman (1) in an experimental approach to the dynamics of rumor. They carried on demonstrations with audiences, six or seven members of which were selected as subjects and asked to leave the room. A picture in slide form was projected on a screen and a description of it given to the first subject, who had been called back into the room, but placed so that he could not see the screen. The first subject passed on a description to the second subject brought into the room. The procedure was repeated under the same conditions until the last subject reported what he had heard.

In analyzing their data, Allport and Postman noted three main processes. One was leveling, or the elimination of details. The second was sharpening "selective perception, retention, and reporting of a limited number of details from a larger context". Leveling and sharpening are reciprocal processes and occur together. The third was assimilation, the modification of a report due to "the intellectual and

emotional context" of the individual's mind. Special interests, one's national, racial, and religious affiliations, etc., all affect the perception and transmission of a report.

As indicated here, descriptions of pictures, objects, or episodes have been used in an attempt to refine the effectiveness of legal testimony (21), predict journalistic aptitude (11), explore some of the phenomena of human learning (2), and investigate the dynamics of rumor perception and transmission (1). The research design described on the following pages also made use of pictorial stimuli, with additional similarities to some of the earlier work in that both a free report and a form of interrogatory were used as methods for eliciting responses. The investigation to be reported here and those already reviewed differ with regard to the purpose of the research, categorization of responses, and analysis and interpretation of the data.

The Problem

To develop a predictor of reporting ability, the assumption had to be made that differential degrees of reporting ability exist among the individuals who are used as test subjects.

This research is concerned with the construction and validation of a test of reporting ability as an aid in selecting product testing panels. If subjects could be preselected by a short, easily administered instrument which would make significant discriminations between those

individuals who offer the most valuable comments about an item tested, and those individuals whose comments are only occasionally helpful, then the product testing process would be refined beyond its present state.

A high level of reporting ability may be considered to consist of two critical components: (1) the ability to attend selectively to those aspects of a situation which are of optimal importance, i.e., ability to observe and (2) the ability to verbalize observations made in this way. Reporting ability is dependent upon both. However, some individuals may be keen observers but have difficulty in communicating their observations. Such people are valuable as test subjects when their observations can be recorded by a recognition test, such as multiple choice or true-false methods. Part of this study is designed to discover the degree to which observing ability and reporting ability are separable. The method used to select reporters was modified to select observers.

With some modification, a reporting ability test could be used to select individuals for any kind of task where the coupling of keen observational ability with adequate linguistic ability was important. Considerations of time and facilities prompted the use of a product testing panel as a criterion group, with the thought that should the test successfully predict good reporters in this group, later validation studies would be devoted to adapting it to other situations.

The ideal panel subject combines several traits and abilities which are useful in product testing. Only a limited attempt will be made to investigate some of these factors. The primary purpose of this research is limited to selecting good reporters.

There are other limitations in this research. Product testing subjects who have been selected by a reporting ability test may not yield opinions which represent those of the population from which they were drawn. Preselected subjects, however, should be able to supply more adequate information that will be helpful in the comparative evaluation of items.

CHAPTER II

DEVELOPMENT OF THE PREDICTOR

The Stimulus Pictures

The stimuli used in this study consisted of a number of pictures which were exposed to the subject one at a time for a fixed time interval. Following the presentation of the stimulus, the subject reported what he had seen in the picture, and took a recognition test on the content of the picture.

Four criteria were established for the selection of the stimulus pictures: first, the picture would contain a large amount of detail; second, the picture would be coherent, have one theme or tell one story; third, the picture would contain some possibilities for ambiguous interpretation; and fourth, a variety of types would be selected, e. g., indoor, sports, personal violence, accidents, military activities, etc.

With these criteria in mind, a search was made for appropriate stimuli in the Library of Congress, in Life Magazine files, in newspapers, and in textbooks. Also available was the photo print library of Guild Photographers, Washington, D. C. From an initial group of over two thousand potential stimuli, sixty were found which met the criteria noted above. These sixty were subjected to further judges' ratings, and the number reduced to twenty. Additional

TABLE I

DESCRIPTION OF THE THIRTEEN PICTURES
SELECTED FOR THE INITIAL PHASE OF THE STUDY

*1. Line drawing of the interior of a street car. Focal point is a Negro and White man, both standing. Other passengers are seated.

2. Line drawing of a battle scene. A Negro soldier is standing, apparently preparing to throw a grenade.

3. Line drawing of a street scene. Prominent objects are a fruit peddler and cart; a boy is running by, grabbing for some fruit.

4. Line drawing of a collision between a street car and an automobile.

**5. Photograph of an intersection in a busy downtown area. Part of the street is being repaired.

***6. Photograph of a room in a slum tenement. A woman is ironing clothes. A number of objects are scattered about the room.

7. Photograph of a backyard slum in Washington, D. C.

8. Photograph of a wash drawing. Cars lined up at a ferry landing.

9. Photograph of a wash drawing. A photographer is taking pictures of a wedding scene.

10. Photograph of a wash drawing. A man painting crosswalk lines on an intersection. Spectators.

11. Photograph of two men picketing for repeal of the amusement tax in front of the White House.

12. Photograph of a cartoon: Primitive men working around a crude bone crushing device.

13. Photograph of a crew repairing telephone lines.

All pictures are black and white.

---

*Pictures 1-4 are from Allport and Postman (1), pp. 71, 67, 70, and 54 respectively.

**Pictures 5 and 13 curtesy Potomac Electric Power Co., Washington, D. C.

***Pictures 6 and 7 curtesy National Capitol Housing Authority, Washington, D. C.

screenings eliminated seven more. The remaining thirteen were used in the pretest of the instrument.

Four copies were made of each picture. These were all 8" x 10" black and white mat prints, singleweight paper, and were mounted on 9" x 11" grey mounting board. A brief description of the thirteen pictures can be found in Table I. Copies of the six pictures used in the final form of the test are included in Appendix A.

The Recall Pretest

The aims of the recall pretest were first, to determine in part which of the pictures were most discriminating; second, to find the optimum length of time for a subject to be exposed to a stimulus to ensure greatest variability of response; and third, to aid in developing a recognition form of the test based on responses furnished by the recall pretest subjects.

Procedure. Seventeen male college students were used as subjects. Each was tested individually. Average time for one subject was about 50 minutes. The testing room contained a table and two chairs, stop watches, a Gray Audograph with conference type microphone, and the stimulus pictures, mounted, and placed face down in a stack on the table. The subject was told that when the test started he was to pick the top picture off the stack, turn it over, and study it intently for a given number of seconds. The pictures were exposed for either 30, 45, or 60 seconds apiece. The

subject was told to stop at the end of the exposure period. Then he reported, as exactly as possible, all that he could remember of the picture. Specific verbal directions given to the subject appear in Appendix B, the Test Administrator's Manual.

The order of presentation of the pictures was rotated for each subject. In this way, each picture appeared in each possible position at least once.

While the subject was reporting what he had seen into the microphone, the experimenter was operating two stop watches. One watch recorded the total time that the subject was reporting. The other watch was of the time-out variety. The experimenter held this latter watch in his hand, and recorded as accurately as possible the time when the subject was actually talking. It was believed that the time factor might be a useful index of reporting ability. Two time measures were available on each subject as a result of the above procedure. One was the total time for each report. The other was the ratio of time actually speaking to total time.

Results. Two criterion analysis of variance designs were set up as tests of significance for the time relationships noted above. Subjects and pictures were the sources of variability in each case. Similar results were obtained for both total time, and the ratio of speaking time to total time; namely, differences among pictures were significantly greater than differences among subjects. This indicated that the time factor was more closely related to the picture than to the

individual. In addition, a Pearson r for total time with ratio total time/speaking time was computed at 0.96. On the basis of these results, the two measures were dropped from further consideration as useful indices of reporting ability.

Of the three stimulus exposure times (30, 45, and 60 seconds) the optimum appeared to be 45 seconds. Notes on the subjects' behavior during the exposure periods suggested that the subjects lost interest in the task towards the end of the 60 second period. Many times they looked up from a picture before the 60 seconds had expired, indicating they were ready to report. This behavior was not manifest during either of the other two exposure periods. Moreover, the 30 second exposure period often brought complaints of not enough time.

From an examination of the subjects' reports, it was apparent that picture #12, the primitive men cartoon, was not discriminating and could be discarded. Judgment on the discrimination characteristics of other pictures could not be made by inspection, and was reserved until after another set of tests on the recognition pretest was done.

The Recognition Pretest

The transcriptions of the verbal reports were used to construct a recognition test which would tap the observational ability of the subject, as distinct from reporting ability. A total of 290 true-false items were constructed, based mainly on the recall reports of the subjects. Since

these items were identified with the indices of reporting ability at the time they were entered into the recognition test, a discussion of these indices follows.

Indices of Reporting Ability. In searching for a predictor of reporting ability, no operational definition of reporting ability was found. It was agreed that a set of hypotheses which could be built into the test, and then separately identified in the form of part scores, would help in isolating the components of the trait to be measured. An attempt was made to determine different scorable indices which might contribute to reporting ability. The best of these comprised the original set of indices and are included in Table II.

Four of the indices in Table II survived, and are keyed into the recognition test at present. These are detail (#1), ambiguities (#3), clothing detail (#4), and inference (#5). The three indices that were discarded are important detail (#2), time factor (#6), and confidence of subject (#7).

Important detail (#2) was discarded because no suitable means of deciding what is important could be found. Judges' ratings of the pictures were not reliable. Also a frequency count of what the subjects themselves recalled, and the order in which they were recalled was affected too much by the disproportionate appearance of a novel object, quite apart from its importance in the picture. The time factor (#6) was discarded for reasons already stated. The subject's indication of confidence in his report (#7) was considered not feasible

TABLE II

ORIGINAL SET OF INDICES OF REPORTING ABILITY

1. Detail - A frequency tabulation of the amount of detail included in the report.

2. Important Detail - Important detail which the subject omits.

3. Ambiguities - Modification of existing detail or addition of detail not in the original stimulus.

4. Clothing Detail - A separate frequency count of clothing and personal item detail.

5. Inference - Reporting as seen, something which was actually inferred.

6. Time Factor - Either total time needed to report, or ratio of speaking time to total time.

7. Confidence of Subject - Subject's estimate of amount of confidence he places in each report in rating scale form.

because in the validation study the confidence rating after each picture would take up too much additional time.

A list of reasons for scoring an item in one category rather than another was compiled for the recognition test. For detail (Table II, #1), an item was scored as true if the object was in the picture, false if the object was not in the picture. An ambiguity (Table II, #3) was keyed for errors of comission on the assumption that the better the observer, the fewer such shifts. Inference (Table II, #5), was scored for an item based on limited or non-existent evidence, usually an individual's trade, profession, race, nationality, religious affiliation, destination, speed; or for the time of day, season of year, or weather. For example,

a woman walking with a child is not necessarily the child's mother. Besides the categories of True and False, there was a category of Not Sufficient Evidence (NSE) for such recognition items.

Procedure. The administration and analysis of the recognition test constituted the final pre-validation phase in the development of the instrument. The purpose of this step was two-fold: first, to find out which of the recognition items were the best discriminators, and second, further to eliminate stimulus pictures, if warranted.

The pictures were arbitrarily divided into two groups of six. Slide transparencies of the pictures, 35mm size, were used in group administrations of the test. The test was given to two class sections of Psychology I, Introduction to Psychology, at the University of Maryland. Each section had an enrollment of about 150 students. One set of six pictures was exposed to each section. Each student recorded his responses on IBM answer sheets, and upon completion of the test, was instructed to indicate how well he could see the screen and to identify his seat number. Even though the projected slide was approximately 48" x 60", and the room equipped with black-out shades, some students sitting at the sides or back reported difficulty in seeing the pictures. Their papers were not used in later analyses.

Since with the student's groups the primary interest was in their responses only, exposure time was kept only approximately constant, and varied from 50 to 55 seconds for

each picture. An additional device that was tried out was a short training period. A specimen picture was projected onto the screen before the first of the test pictures. A number of sample questions, inclusive of all the response categories, were read orally. The group answered the sample items aloud. During the training period and the test proper, there seemed to be no ambiguity with respect to the use of each of the response categories (true, false, not sufficient evidence, don't know). Motivation appeared to be high throughout. The novelty of the task appealed to the subjects.

Results. In preparing to item analyze the responses to the recognition test using the two psychology classes as subjects, certain answer sheets were excluded. First, answer sheets having over 20% "don't know" responses were eliminated. Second, the papers of those who stated inability to see the projection screen clearly were eliminated. Third, students who sat in seats other than the center section, first ten rows, were eliminated. After these exclusions there remained 74 answer sheets from one section, and 62 from the other. The answer sheets were scored by the formula: right minus half the number wrong, since there were three predictable alternatives for each answer. The scored answer sheets were then used to identify the upper 27% and lower 27% of the subjects in each class section. Using the answer sheets in these upper and lower groups, an item count was made on the IBM graphic item counter. As

an indication of how well each item discriminated between the high and low scores, a biserial correlation coefficient was assigned to each item*. The range of correlation coefficients so obtained was -.04 to +.74 for both groups of subjects. Items with correlation coefficients of less than .40 were dropped. The remaining items were the ones used as a basis for all other analyses, and comprised the final form of the recognition test for those pictures which were included. (See Appendix C4 for final form).

The pictures were further analyzed to identify those that might still be discarded. A picture on which subjects showed wide variability of response was assumed to be a more useful discriminator than one where there was little variability. Accordingly, using the total sample, a standard deviation was computed for each picture score (based on the recognition test) and was one criterion of whether the picture should be retained. Since all the pictures were presumably measuring the same set of factors, a correlational analysis was also used to help determine which to retain. If two pictures correlated highly, for example, one of them might be discarded, other things being equal, since they were both measuring the same thing. Data on the variability of the pictures are presented in Table III.

---

*To facilitate assignment of the biserial coefficients, the table was used which appears on pp. 347 ff. of Thorndike, Robert L., Personnel Selection, New York: John Wiley and Sons, 1949.

TABLE III

MEAN NUMBER OF CORRECT ITEMS AND STANDARD DEVIATIONS FOR EACH OF THE PICTURES

Set I, N=74

	All Pictures	Picture* 1	3	5	7	8	10
Mean	45.01	8.11	7.03	6.13	8.28	8.19	7.01
S.D.	9.34	4.54	2.14	1.62	2.13	1.93	2.24

Set II, N=62

	All Pictures	Picture 2	4	6	9	11	13
Mean	42.11	4.53	8.73	9.58	9.47	5.85	2.63
S.D.	8.82	1.86	3.05	2.24	2.64	1.76	0.87

*The numbers which identify the pictures correspond to the numbers in Table I, page 12.

Table III is partially misleading unless one considers the variation in the number of items per picture. There is a range of from 6 items on picture #13, to 17 items on pictures #1, #4, and #9. With the possible exception of picture #1 and picture #4, the variabilities shown by the standard deviations are too homogeneous to permit satisfactory screening on the sole criterion of within picture variability.

In some cases intercorrelations of the scores of the pictures in each set provided another means for differentiating between the more and less useful pictures. However, it was not possible to intercorrelate pictures from Set I to Set II because a different sample was used for each set. But the correlation coefficients are homogeneous enough in

both sets so that it would be safe to speculate that were all twelve pictures administered to one sample and the resulting scores intercorrelated, the matrix thus produced would not contain a range of coefficients much different from either of the samples shown in Table IV. This opinion gains credence from having the same types of pictures appear in each set.

Table IV shows that the highest correlation coefficients were .54 for #7:#8, and .53 for #7:#5 in Set I; and .53 for #2:#13 in Set II. In no case were the correlation coefficients high enough to warrant omitting a picture because another picture was measuring the same components.

TABLE IV

ZERO ORDER CORRELATION COEFFICIENTS BETWEEN THE RECOGNITION TEST SCORE FOR EACH PICTURE WITH EVERY OTHER PICTURE IN ITS SET

Set I, N-74

Picture*	#1	#3	#5	#7	#8	#10
#1	-	.23	.33	.30	.23	.23
#3		-	.42	.28	.13	.43
#5			-	.49	.53	.24
#7				-	.54	.17
#8					-	-.17

Set II, N-62

Picture*	#2	#4	#6	#9	#11	#13
#2	-	.29	.44	.35	.14	.53
#4		-	.44	.36	.37	.13
#6			-	.45	.27	.37
#9				-	.34	.20
#11					-	.00

*The numbers which identify the pictures correspond to the numbers in Table I, page 12.

The pictures which were finally selected for inclusion in the final form of the test were #1, #4, #6, #8, #9, and #11. Two considerations entered into this choice besides the data in Tables III and IV. These were that the pictures selected were the more discriminating ones in the pretest and that more detail and variety of types of response were to be found among the selected six. The decision to use only six pictures in the final form was a result of experience in the pretest when all 13 pictures were used. Economy of time, and maintenance of motivation needed to be insured, and six pictures were optimum for both.

For convenience, the identifying symbol was changed for each of the six terminally chosen pictures. Upper case letters were assigned the pictures in this order: picture #1 became A, #11-B, #8-C, #4-D, #6-E, and #9-F. In the remainder of this dissertation, and in the appendices, the pictures will be identified by letter.

Correlation of Picture Score with American Council on Education Psychological Examination. Scores on the American Council on Education Psychological Examination, College Edition, 1948, administered routinely to entering freshmen at the University of Maryland, were available for thirty-eight of the subjects in each of the groups. There are three scores for each individual: a subscore on a linguistic factor, a subscore on a quantitative factor, and a total weighted score. Correlation coefficients were computed between both of the subscores and the total score on the

recognition test. For Set I, the recognition test and linguistic factor correlated .26; the recognition test and quantitative factor correlation .29. For Set II, correlation with the linguistic factor was .17, with the quantitative score .72. With the exception of the last result, we are able to say for the samples used, that there seems to be no significant relationships between recognition test score and A.C.E. score. Since the A.C.E., especially the linguistic factor, is fairly highly correlated with other group tests of mental ability, we may conclude that for the two samples studied the recognition test is not measuring what is commonly called intelligence.

Analysis of the Indices of Reporting Ability. The recognition test items were keyed so that they could be scored in still another way. Each of the pictures contained items which are assignable to one of the four categories: detail, clothing detail, ambiguities, and inference. These four comprised the indices of reporting ability before validation.

The pictures in Set I contained 45 detail items, 21 ambiguity, 17 clothing detail, and 13 inference. In Set II, there were 46 detail items, 17 ambiguity, 15 clothing detail, and 12 inference. All categories are mutually exclusive except clothing, which is scored as part of another category as well as by itself.

Correlational analyses for each of the categories for each set indicate some interesting relationships. These data are shown in Table V.

TABLE V

ZERO ORDER CORRELATION COEFFICIENTS AMONG THE FOUR INDICES OF REPORTING ABILITY

Categories	Set I (N-74)	Set II (N-52)
Ambiguity (#3)/detail (#1)	.61	.36
Ambiguity (#3)/clothing detail(#4)	.60	.54
Ambiguity (#3)/inference (#5)	-.04	.15
Detail (#1)/clothing detail (#4)	.73	.74
Detail (#1)/inference (#5)	.07	.17
Clothing detail(#4)/inference(#5)	.17	.17

With one exception, the highest correlations are between clothing detail (#4) and detail (#1), and clothing detail (#4) and ambiguity (#3). Consistently low correlations are associated with the inference index (#5). It seems fairly certain that the inference index (#5) is measuring something different from the other three categories. The array of coefficients suggests that all four indices be retained and examined more intensively as part of the validation process. These results are comparable for Sets I and II, indicating high reliability of the data.

Summary of Pretest Results

Pretest results permitted modifications in the original design of the test with respect to the following points:

1. An initially large number of stimulus pictures were screened down to those which were most discriminating.

2. Steps in the process, such as the length of stimulus exposure time, and recording of results, were standardized for ease and expediency of administration.

3. Through an item analysis procedure the original form of the recognition test consisting of 290 items was reduced to a final form consisting of 89 items.

4. Four of the seven proposed indices of reporting ability were accepted as feasible.

5. A scoring system was established for both the verbal report and the recognition portion of the instrument.

CHAPTER III

THE VALIDATION STUDY

The Subjects

The data for the validation study were collected at Fort Lee, Virginia, by the personnel of the survey division, Quartermaster Board. One of the missions of the survey division is to conduct acceptance tests on new articles of clothing and personal equipment, as a factor in deciding which of them should be made part of general issue. Data are often collected by the panel method. Panel members are officers and enlisted men stationed at Fort Lee or other posts, whose routine activities prescribe articles of the type being tested.

The only acceptance testing panel for whom records for at least one complete survey were available, and who were themselves available at Fort Lee, were part of the group of officers who participated in the "cloth summerweight" uniform study (QMB s-264) during 1951. In the course of this study, the panel members alternated wearing two uniforms, identical except for the composition of the fabric. Uniforms were identified only by letter. Periodic reports in the form of interviews and questionnaires were required of the subjects.

Of the officers who took part in this survey, 55 were stationed at Fort Lee in March, 1952, and were made available

as subjects for the validation phase of the reporting ability test. Part of this group consisted of individuals who are instructors in Quartermaster Corps school at Fort Lee. Pertinent biographical data on all subjects may be found in Appendix D.

Purpose of Validation Study

The validation study was designed to yield information on the following problems:

1. The degree to which responses on both the recall and recognition portions of the reporting ability test predict an independent criterion of reporting ability.
2. The contention that instructors at QMC school (an "expert group") are better acceptance testing subjects than others without similar training and experience.
3. The relative importance of each of the indices of reporting ability.
4. The relationship of selected biographical data to reporting ability. In this study, these included rank, time in service, education, as well as assignment as instructor in QMC school.
5. The degree to which the recall and recognition portions of the reporting ability test measure the same thing.

Selection of a Criterion

Three different criteria of reporting ability were available for the validation of the test. Only one was used.

The first criterion that was proposed was a set of responses to open ended questions obtained by questionnaire and interview. These were part of the questions which the subjects

answered in the course of the "cloth summerweight" study. It was believed that these responses could be scored for information content by personnel at the Quartermaster Board, and the resulting scores correlated with the scores on the reporting ability test. Less than half of the subjects offered comments of more than a few words to any of the open-ended questions. For all the questions together, there were still a number of "no comment" responses. Data would not have been available for each subject, and these answers to questions were dropped from further consideration as a source of criterion data.

Two kinds of ratings of reporting ability were obtained. One set of ratings was made by four of the personnel at QMB who had had contact with the subjects during the "cloth summerweight" study. However, since seventeen subjects were not known well enough to be rated by any of the judges, this measure was also dropped from consideration as a possible criterion.

The criterion used was a set of ratings made by each of the subjects. After the administration of the reporting ability test, each subject was handed a mimeographed sheet containing the names of all the subjects in the validation study (who were also panel members for the "cloth summerweight" survey). Each subject was first asked to place a check mark beside the name of each person on the list with whom he was acquainted. Then he was asked to rate as many of the checked names as possible on the continuum of reporting

ability. Exactly what was desired was carefully explained. The subject placed a plus sign beside all those who were as good or better than he was as a reporter, and a minus sign beside all those who were inferior to him in reporting ability. The subject only rated as many of the checked names as he wished. Detailed instructions for the rating procedure may be found in the test administrator's manual, Appendix B.

Each subject was scored for the number of recognitions (checks before his name), and number of plus and minus ratings. Ten subjects received fewer than ten recognitions and were dropped from the validation group. The criterion score for the remaining subjects consisted of the algebraic sum of the plus and minus ratings, plus fifteen. Adding fifteen to each score insured all subjects of receiving a positive criterion score. The raw rating data and the coded scores may be found in Appendix E.

Validation Procedure

The procedure for validation is outlined in detail in the administrator's manual, Appendix B. The test was administered individually by three enlisted men at the Quartermaster Board. Each of the administrators had a baccalaureate degree in psychology, and had had extensive interviewing experience at the Board. Before meeting with any of the subjects, they participated in a special six hour training period on the procedure to be followed.

Briefly, the procedure in each test session was as follows: each subject made a verbal report on each of the six pictures, which was recorded for later transcription. A recognition test on the same picture followed each verbal report. On completion of this protocol for all six pictures, the subjects completed a rating scale form, and filled out a biographical inventory. After dismissing the subject, the test administrator recorded any pertinent comments on the conduct of the interview.

Results

Scoring of Responses. The verbal reports were transcribed verbatim, typed double spaced, with no new picture recall starting on the same page where one had been completed. One carbon copy was made, to be used later for determining scoring reliability.

Each report was scored for the four indices of reporting ability: detail, clothing detail, inference, and ambiguity. In addition, an estimate of the length of each report was made by counting the number of lines of copy. In several instances it was necessary to make an arbitrary decision for scoring some part of a report. These were noted and adhered to in scoring all other reports on the same picture. A copy of the special notes for each picture may be found in Appendix C2. The general instructions for scoring reports are included in Appendix C1. A sample scored report for each picture is available in Appendix C3.

The recognition test answers were scored by the formula "rights only" for each of the four indices of reporting ability. In addition, the number of "don't know" responses was tabulated and included as another index. An answer was tabulated as "don't know" either when the subject had checked that category, or when the answer space had been left blank. There were few blank spaces in all the question booklets. A copy of the question booklet used in the recognition test is in Appendix C4. The scoring key for the recognition test is included in Appendix C5.

The recall and recognition parts of the test had four shared indices: detail, clothing detail, inference, and ambiguity. Each part had one possible index independent of the other. For the recall responses, this was an estimate of the length of report. For the recognition responses, it was the number of "don't know" answers. The scores on the indices for each subject for all six pictures were the data for analysis. These raw data may be found in Appendix F1 and F3 for recall and recognition, respectively. In addition, the raw scores for each picture with indices combined were used for other analyses, and appear in Appendix F2 and F4 for recall and recognition, respectively. The method of scoring the criterion ratings has already been discussed (see pages 29-30). The biographical data were summarized for later use and are included in Appendix D.

Reliability of Test. The reliability of this test is dependent upon two kinds of measures: one is a measure of

the internal consistency of the instrument, the other a measure of the agreement among scorers on the recall portion of the test. An estimate of the internal consistency of the instrument was made by dividing the six pictures into two groups by use of a table of random numbers, scoring each group separately, and computing a correlation coefficient for the two sets of scores obtained. In the case of the recognition test, one half of the pictures contained one more item than the other half. This item was eliminated by use of a table of random numbers. Uncorrected split half reliability coefficients were found to be .84 for the recall portion of the test, and .50 for the recognition part, both based on 55 cases. These results indicate that the recall part of the test is satisfactorily stable, but the recognition test is not stable enough for further use in its present form.

All 330 of the recall reports (55 subjects, 6 pictures each) were scored by the same individual. The carbon copies of 48 of these reports, eight for each picture, were divided equally among four individuals, who scored them using only the instructions for scorers and notes on each picture as guides. The scores on each report for detail, clothing detail, inference, and ambiguity were summed. The sums so obtained were combined for pictures A, B, and C, and for D, E, and F, and correlated with similarly obtained scores from the original set of scored reports. The correlation

coefficient between rescored reports and originally scored reports was .92.

Validity of Test. The extent to which the recall and recognition portions of the test predicted judges' ratings was estimated by computing two multiple correlations based on the indices of reporting ability, one for recall scores and the other for recognition scores. The number of cases in the analyses in both instances was 45, the ten subjects with the fewest number of judges' recognitions on the rating scale having been dropped from the population. (See page 30).

The separate scores for the proposed indices of reporting ability and the criterion scores were intercorrelated to produce matrices of zero order correlations. The matrix of intercorrelations for recall scores is shown in Table VI, for recognition scores in Table VII.

TABLE VI

ZERO ORDER CORRELATIONS BETWEEN EACH INDEX OF REPORTING ABILITY AND EVERY OTHER INDEX (BASED ON RECALL SCORES), AND WITH THE CRITERION OF JUDGES' RATINGS

	*Index				
	**D	C	I	A	L	Criterion
D	-	.829	.092	.572	.815	.094
C		-	.299	.500	.750	-.012
I			-	.251	.541	-.364
A				-	.598	-.203
L					-	-.095

*The letters are abbreviations of the following: D-detail, C-clothing detail, I-inference, A-ambiguity, L-length of report in lines.

**The detail score includes the clothing detail score.

TABLE VII

ZERO ORDER CORRELATIONS BETWEEN EACH INDEX OF REPORTING ABILITY AND EVERY OTHER INDEX (BASED ON RECOGNITION SCORES), AND WITH THE CRITERION OF JUDGES' RATINGS

	*Index				
**D	C	I	A	DK	Criterion
D -	.426	.302	.217	.254	.265
C	-	.211	.150	.174	.181
I		-	.114	.143	.134
A			-	.110	.103
DK				-	.130

*The letters are abbreviations of the following: D-detail, C-clothing detail, I-inference, A-ambiguity, DK-"don't know" responses.

**The detail score includes the clothing detail score.

A comparison of Tables VI and VII indicates some interesting differences among the intercorrelations of the indices on the recall (Table VI) and recognition (Table VII) portions of the test. The range of correlations is markedly greater for the recall matrix (-.364 to .829) than for the recognition matrix (.103 to .426). This tendency for more homogeneous correlations to appear in the recognition matrix is further shown in comparing the correlations for each index with the criterion: for recall they vary from -.364 to .094 and for recognition they vary from .103 to .265. All of the correlations in the recognition matrix are positive; four of the correlations in the recall matrix are negative, all of them including the criterion as one variable.

The Doolittle method was used to compute the Beta coefficients for both the recall and recognition solutions for the multiple correlation, corresponding to the model in Peters and Van Voorhis text (16). Multiple correlation coefficients of .491 and .251 were computed for the recall and recognition matrices respectively. A test of significance was made on both of these coefficients. The probability of a five variable multiple correlation coefficient of .491, based on 45 cases, occurring through chance is less than .05 but greater than .01. The probability of occurrence of a coefficient of .251, under the same conditions, is greater than .05. From these data, we may conclude that the predictive efficiency of the recall portion of the test is somewhat greater than that of the recognition portion, and advise that the latter be dropped or considerably modified before any further attempts be made to use it as a predictor of reporting ability.

The detail scores for both the recall and recognition portions of the test included the clothing detail scores. Clothing detail was also entered as a separate score. Intercorrelations of the detail score with every other variable were recomputed, subtracting the clothing detail score in each case. The resulting differences are shown in Table VIII and were not considered great enough in either case, to warrant computation of an additional multiple correlation with this new set of intercorrelations.

TABLE VIII

INTERCORRELATIONS OF THE DETAIL INDEX WITH ALL OTHER VARIABLES WHEN CLOTHING DETAIL IS INCLUDED AND EXCLUDED FROM THE DETAIL SCORE

	Recall *Index					
	C	I	A	L	DK	Criterion
Included	.829	.092	.572	.815	-	.094
Excluded	.784	.061	.567	.802	-	.106
	Recognition					
Included	.426	.302	.217	-	.254	.265
Excluded	.417	.301	.217	-	.257	.267

*The letters are abbreviations of the following: C-clothing detail, I-inference, A-ambiguity, L-length of report in lines (recall only), and DK-number of "don't know" responses (recognition only).

Relation of Recall and Recognition Parts of Test. To find the extent to which the recall and recognition portions of the test are measuring the same thing, a correlation coefficient was computed. The raw data used in this computation were the summed scores for the four indices which both parts of the test shared. All 55 cases were used in this computation. The correlation coefficient between recall scores and recognition scores was -.105, and is not significantly greater than zero at the .05 level of probability. An interpretation and discussion of this correlation may be found in Chapter IV.

Contribution of Each of the Indices to the Multiple Correlation. An estimate of the unique contribution of each of the indices to the prediction of the criterion of judges'

ratings may be made by comparing the Beta coefficients for each index. These data were obtained as part of the process of computing multiple correlations from the matrices in Tables VI and VII, and are shown in Table IX.

TABLE IX

BETA COEFFICIENTS FOR EACH INDEX OF REPORTING ABILITY FOR BOTH THE RECALL AND RECOGNITION PORTIONS OF THE TEST

Index	Recall	Recognition
Detail	.226	.197
Clothing detail	-.032	.072
Inference	-.336	.048
Ambiguity	-.297	.038
Length of report	.120	-
N "DK" responses	-	.057

Table IX indicates that inference, ambiguity, and detail are the most useful indices in the recall portion of the test, while only the detail score contributes to the recognition score. The initial hypotheses with respect to the inference and ambiguity indices seem to have been substantiated in the recall portion of the test. The hypotheses were that both of these indices are inversely related to reporting ability.

Contribution of Scores Obtained From Each Picture to Total Score. An analysis of recall scores by picture was made by summing the index scores for each picture. The indices combined were detail (excluding clothing detail), clothing detail, inference, and ambiguity. The total score for each subject on all six pictures was used as a criterion, and a

matrix of intercorrelations of the pictures with each other and with total score was obtained. All 55 cases were used. The Doolittle method was used (16), and Beta coefficients computed to determine how much of the total score was accounted for by each of the six pictures. Under these circumstances, with total score as the criterion, the multiple correlation should account for all of the variability in the criterion. The multiple correlation coefficient computed from the Beta coefficients was .993. The matrix of zero order correlations is shown in Table X.

TABLE X

ZERO ORDER CORRELATION COEFFICIENTS FOR EACH PICTURE WITH EVERY OTHER PICTURE AND WITH THE CRITERION OF TOTAL SCORE, FOR THE RECALL PORTION OF THE TEST

	Picture						
	A	B	C	D	E	F	Criterion
A	-	.707	.499	.642	.574	.723	.821
B		-	.729	.720	.653	.661	.885
C			-	.593	.610	.672	.814
D				-	.675	.646	.834
E					-	.703	.834
F						-	.872

The range of correlation coefficients in Table X is from .499 (A with C) to .885 (B with criterion). None of the pictures is as highly correlated with any other picture as it is with the criterion. The correlations of the pictures with the criterion are notably homogeneous. The Beta coefficients computed from this matrix are shown in Table XI.

TABLE XI

BETA COEFFICIENT FOR EACH PICTURE WITH THE CRITERION OF TOTAL SCORE FOR THE RECALL PORTION OF THE TEST

Picture	Beta Coefficient
A	.187
B	.206
C	.204
D	.167
E	.205
F	.199

Table XI shows that each picture contributed approximately the same to total score. No picture was more valuable than any other in the test series.

Analysis of the Biographical Data. In order to determine the relationship between reporting ability and selected biographical data, chi-square analyses were made using both recall scores on the reporting ability test, and judges' ratings. A separate analysis was made for each of these variables: rank, time in service, education, and whether or not the subject was currently assigned as an instructor in QMC school. For the purposes of analysis by recall score, the scores were divided into two class intervals, encompassing the range of total scores, and splitting the distribution into approximately equal parts. The sub-ranges were 74-165, and 166-277. The coded scores for judges' ratings were also divided into two class intervals, with about half the coded scores in each interval: sub-ranges of 9-22, and 23-42.

The biographical data for three of the variables were dichotomized as follows: for rank, company grade (including one warrant officer), and field grade officers; time in service to nearest year, 1-10 years and 11-25 years. Either "yes" or "no" were the class intervals for current assignment as QMC school instructor. The education variable was trichotomized into these categories: high school, some college but no degree, college degree granted.

Tables XII and XIII show the results from the chi-square analyses for each of the biographical variables, using recall scores and judges' ratings respectively. Table XII is based on 54 cases. Table XIII is based on 44 cases (ten subjects who were dropped from the criterion group).

Table XII (recall scores) shows no significant chi-squares. Table XIII (judges' ratings) shows the rank variable associated with a chi-square significant at the .01 level.

TABLE XII

CHI-SQUARE, DEGREES OF FREEDOM, AND PROBABILITY OF OCCURRENCE OF CHI-SQUARE BASED ON RECALL SCORES

Variable	*Chi-square	DF	p
Rank	0.95	1	>.05
Time in service	2.67	1	>.05
Education	1.72	2	>.05
QMC instructor	0.37	1	>.05

*Corrected for continuity in cells with observed frequencies under ten.

TABLE XIII

CHI-SQUARE, DEGREES OF FREEDOM, AND PROBABILITY OF OCCURRENCE OF CHI-SQUARE BASED ON CODED CRITERION SCORES

Variable	*Chi-square	DF	p
Rank	9.00	1	<.01
Time in service	0.00	1	>.05
Education	1.42	2	>.05
QMC instructor	0.14	1	>.05

*Corrected for continuity in all cells: (N equals 44).

None of the other variables in Table XIII has a significant chi-square.

That the rank variable has a significant chi-square when judges' ratings are used as a variable (Table XIII), but not when recall scores on the reporting ability test are used as a variable (Table XII), is a reflection of a source of error in the criterion scores. Regardless of how much the anonymity aspect of the rating procedure was stressed to the subject, there probably was a tendency to accord deference to higher ranking officers. There is probably also a positive halo effect related to higher rank. Also important is the influence and awareness in officers of ratings which are part of efficiency reports, periodically filled out by an officer for subordinate officers in his command. These efficiency reports are a factor in determining promotion. Since the subjects for this study give and/or receive efficiency reports based on rating devices, it is likely that their attitudes toward efficiency

ratings, built up over a period of years, may have become generalized to any kind of task where they are required to rate fellow officers.

CHAPTER IV

DISCUSSION OF RESULTS

Summary of Results

Analysis of the data obtained from the validation study yielded the following results:

1. Uncorrected split half reliability for the recall portion of the test was .84. For the recognition portion, .50. Reliability of the scoring procedure on the recall portion of the test as estimated by correlating reports scored independently by four individuals with reports scored by the author, was .92.

2. A multiple correlation coefficient of the scores on the instrument with the criterion of judges' ratings was .49 for the recall part of the test, .25 for the recognition part.

3. Beta coefficients computed for the indices of the reporting ability with the criterion of judges' ratings were -.336, -.297, and .226 for inference, ambiguity, and detail respectively, on the recall part of the test. For recognition, a Beta coefficient of .197 was obtained for detail. The Beta coefficients for the other proposed indices on either sub-test were about zero.

4. A set of Beta coefficients for the six pictures against the criterion of total score were computed for the recall sub-test. These Beta coefficients indicate that all six pictures contribute about equally to predicting the total score.

5. A correlation coefficient computed for scores on the recall part of the test with scores on the recognition part of the test was -.105.

6. A chi-square analysis of four biographical variables: rank, time in service, education, and assignment as QMC school instructor, yielded no significant chi-squares when recall scores were used as a basis for entering frequencies. When the criterion scores were used, the rank variable alone yielded a significant chi-square, and that at the .01 level of significance.

Discussion of Results

Of the two parts of the reporting ability test, recall and recognition, the recall portion has emerged as markedly superior of the two, with respect to both reliability and validity (see page 44, #1, #2). Any further application of this test should be limited to obtaining responses by verbal report; the recognition test does not yield useful data in this situation.

The failure of the recognition test could be ascribed to pretesting it on a sample different from the sample used in the validation study; yet the recall pretest utilized a sample from the same population as was drawn on for the recognition pretest. A more satisfactory reason for the failure of the recognition test lies in the nature of reporting ability itself. The recognition test was constructed to tap observational ability, as distinct from reporting ability in which the verbal factor is prominent. This verbal factor was utilized in the recall part of the instrument. The ability to observe is probably dependent upon the ability to verbalize observations. Just as most thoughts and images cannot be experienced without a verbal

label, so perhaps the ability to observe does not exist, or at any rate is not measurable apart from the ability to report such observations.

The above contention can be substantiated by noting first, that considerable time was spent in refining the items for the recognition test to maximize their discriminability. Second, and of primary importance, the recognition test was administered in each case, immediately after the subject had completed the verbal report on the picture to which he had just been exposed. Logically, we might expect that the subjects who scored high on the recognition test would score high on the recall test. They did not. The items on the recognition in many cases duplicated parts of the subject's verbal report. Yet the correlation between the recall and recognition parts of the test was -.105. Even considering the more limited range of the recognition test (89 items), it is difficult to postulate any reason for the correlation coefficient of -.105 other than that already expressed: observational ability is contained in reporting ability, the reverse is not true.

The criterion of judges' ratings was the best obtainable under the circumstances, but was not satisfactory. A more rigorous criterion of reporting ability might have correlated more highly with the verbal reports on the pictures. Such a criterion might be based on a period of simulated product testing, with the panel unaware that they were reporting ability test subjects until the completion

of the product testing period, when the reporting ability test would be administered. Reports of panel members of the products tested could be scored for content, and these scores used as the criterion of reporting ability.

The usual sources of errors in rating procedure were enhanced in this study, because the raters were army officers. Some evidence to substantiate the above statement is the significant chi-square associated with the rank variable when ratings were used as a basis for entering frequencies, but not when recall scores were similarly used. In addition, the task of asking officers to rate other officers on observing and reporting ability was a novel one, and it is likely that a variety of different sets influenced the rating procedure and so introduced additional error.

On the basis of the analysis of the validation study data, the indices of reporting ability which are the most useful are detail, inference, and ambiguity. The proposed indices of clothing detail, and length of report measured in lines, may be discarded. Of the three indices retained, both inference and ambiguity are associated with negative Beta coefficients. Thus, future scoring devices should be keyed to penalize subjects according to the number of inferences and ambiguities which appear in the report. A system of weighting raw scores could be based on the Beta coefficients.

Since each picture is about as useful as any other in predicting total score, no conclusions may be drawn with respect to eliminating additional pictures. It would be possible, to use only four pictures, instead of all six. Which two are dropped is arbitrary.

The chi-square analysis of biographical data indicates that the factors of rank, time in service, education, and current assignment as QMC school instructor are not correlated with ability to report. Especially interesting is the insignificant chi-square associated with assignment as QMC school instructor. It is the opinion of many of the personnel at the Quartermaster Board that officers who are instructors in Quartermaster school are better product testing subjects than those who have not had this training and experience. In certain aspects of product testing, instructors may be superior. They do not appear different from non-instructors with respect to reporting ability.

The test of reporting ability was validated on a product testing panel. There is no reason why its usefulness could not be extended to selection of subjects in almost any area where the ability to make keen observations coupled with the ability to verbalize those observations is important. It is a wide field, ranging from selection of subjects for psychophysical investigations to selection of competent courtroom witnesses. Whipple, in 1910, wrote (21):

> .....in time there may be developed a satisfactory series of diagnostic tests for estimating the reliability of witnesses.....

That time has not yet arrived. To some degree, this research may be a contribution to that end.

CHAPTER V

SUMMARY AND CONCLUSIONS

The purpose of this research was to construct and validate a test of reporting ability to be used in selecting subjects for product testing panels.

Thirteen stimulus pictures, selected from an initial group of over two thousand, were administered in two pretests to students at the University of Maryland. In the first pretest, the pictures were exposed to each of 17 students for 45 seconds. Following the exposure period, the student reported as exactly as possible all that he remembered of the picture. The verbal reports so obtained were used as a basis for (1) constructing items for a recognition type test, and (2) developing a scoring system for recall data.

A 290 item true-false type recognition test, with items based on subjects' responses in the recall pretest, was administered to 300 students in two sections of Psychology I, Introduction to Psychology, at the University of Maryland. The stimuli were presented in the form of slides projected on to a screen.

Based on information from the two pretests, the instrument was refined in these respects: (1) six of the original thirteen stimulus pictures were found to be most discriminating, and were retained. The other seven were

dropped; (2) 89 of the recognition test items were retained; the others were dropped because they were found to be poor in discriminating between high and low scores on the test; (3) four of seven proposed indices of reporting ability were retained; the other three were dropped because they did not discriminate among subjects, or because they were found to be awkward to score.

The predictor was validated on 55 U. S. Army officers at Fort Lee, Virginia. These officers had all been part of a product testing panel, and their pooled ratings of each others' reporting ability constituted the criterion of reporting ability used in the analysis of the validation study data. Both the recall and the recognition portions of the test were administered individually to the subjects.

Analysis of the validation study data yielded information upon which the following conclusions are drawn:

1. The recall portion of the test is superior to the recognition portion in predicting the criterion. The recognition test should be eliminated from further use as a predictor of reporting ability.

2. The inference, ambiguity, and detail indices of reporting ability should be retained. The indices of clothing detail and length of report should be dropped as poor predictors of the criterion.

3. Four of the pictures would probably predict the criterion as well as six. Which two are dropped is arbitrary.

4. The biographical factors of rank, time in service, education, and assignment as instructor in QMC school are not correlated with reporting ability.

SELECTED BIBLIOGRAPHY

1. Allport, G. W. and Postman, L. The Psychology of Rumor. New York: Henry Holt and Company, 1947.

2. Bartlett, F. C. Remembering. Cambridge: Cambridge University Press, 1932.

3. Berelson, B. and Lazarsfeld, P. F. The analysis of communication content. Bureau of Applied Social Research, Columbia University, 1948.

4. Binet, A. La Suggestibilité. Paris, 1900.

5. Binet, A. Psychologie individuelle. La description d'un objet. A.P., 3, 1897, 296-332.

6. Bruner, J. S. The dimensions of propaganda: German shortwave broadcasts to America. J. abn. and soc. Psychol., 1941, 36, 311-337.

7. Burtt, H. E. Applied Psychology. New York: Prentice-Hall, Inc., 1948.

8. Carmichael, L., Hogan, H. P., and Walter, A. A. An experimental study of the effect of language on the reproduction of visually perceived form. J. exp. Psychol., 1932, 15, 73-86.

9. Cattell, J. Measurements of the accuracy of recollection. Science, N.S., 2, 1895, 761-766.

10. Coombs, C. H. Some hypotheses for the analysis of qualitative variables. Psychol. Rev., 1948, 55, 167-174.

11. Freyd, M. A test series for journalistic aptitude. J. appl. Psychol., 1921, 5, 46-56.

12. Jones, D. B. Quantitative analysis of motion picture content. Publ. Opin. Quart., 1942, 6, 411-428.

13. McGeoch, J. A. The influence of age and sex upon ability to report. Amer. J. Psychol., 1928, 40, 458-466.

14. Meumann, E. The Psychology of Learning. New York: D. Appleton and Company, 1913.

15. Munsterberg, H. On The Witness Stand. New York: 1908.

16. Peters, C. C. and Van Voorhis, W. R. Statistical Procedures and Their Mathematical Bases. New York: McGraw-Hill Book Company, Inc., 1940.

17. Sargent, S. S. and Saenger, G. Analyzing the content of mass media. J. Soc. Issues, 1947, 3, 33-38.

18. Stern, L. W. Zur psychologie d. aussage. Exp. untersuchungen ueber erinnerungstreue. Zeits. f.d. ges. strafrechtwissenschaft, XXII, 1902.

19. Thorndike, R. L. Personnel Selection. New York: John Wiley and Sons, 1949.

20. Whipple, G. M. Manual of Mental and Physical Tests. Baltimore: Warwick and York, 1910.

21. Whipple, G. M. The observer as reporter: A survey of the "psychology of testimony". Psychol. Bull., 1909, 6, 153-170.

22. Woodrow, H. The ability to learn. Psychol. Rev., 1946, 53, 147-158.

APPENDIX A

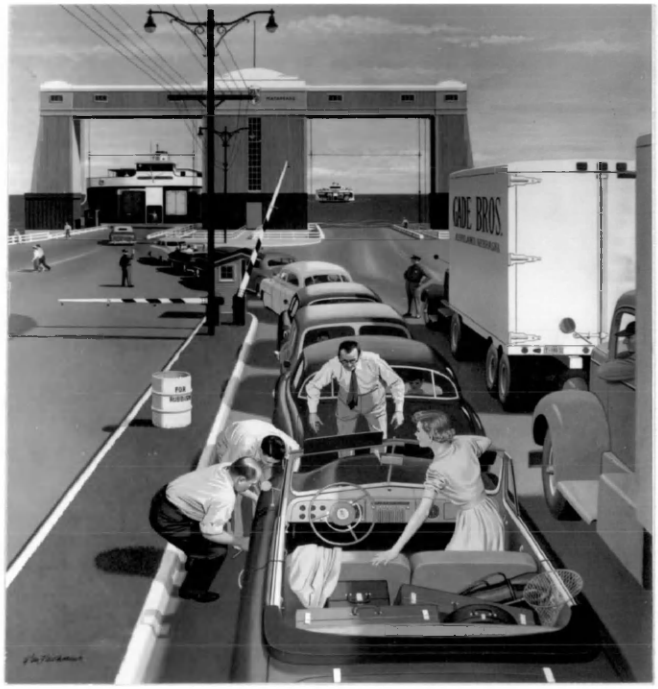
The Six Stimulus Pictures Used in the Final Form of the Test

For administration, each picture was an 8"x10" mat print, mounted on 9"x11" grey mounting board

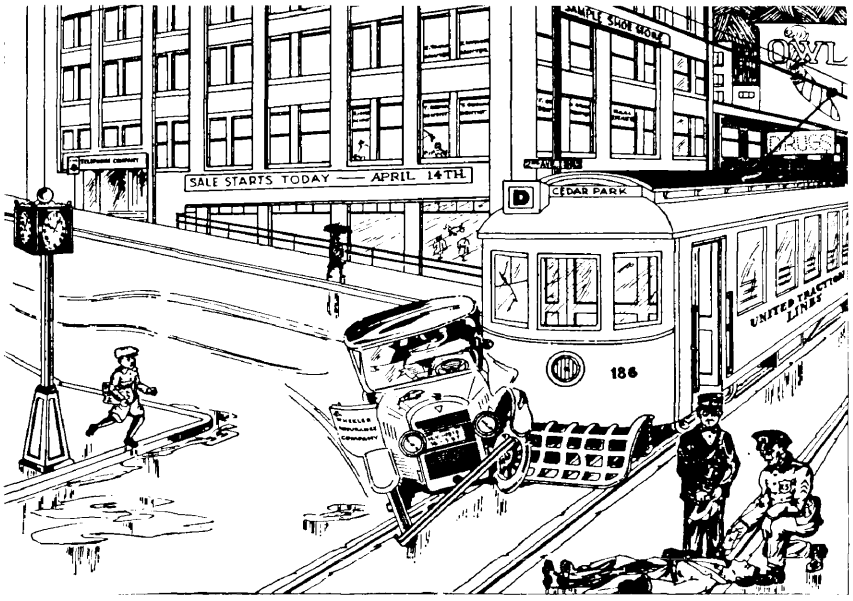

Smoke
LUCKY RAKES
THEY SATISFY
GOSLING SOAP
CHEAPER
BETTER
PURER
CAMP KILEWILO
RESTRICTED CLIENTELE
VE. EXPRESS
AN CORTLAND PK.
DYKEMAN ST.


THE WAR IS OVER
20%
AMUSEMENT TAX
LINGERS ON!
WE DEMAND REPEAL OF THE
20% AMUSEMENT TAX
Congress
U.S.A.

Picture B



Picture C


SALE STARTS TODAY — APRIL 14TH
CEDAR PARK
186
UNITED TRACTION LINES

Picture E

Picture F

APPENDIX D

Test Administrator's Manual

IMPORTANT: THIS BOOKLET SHOULD BE TREATED AS CLASSIFIED MATERIAL DURING THE TESTING PERIOD

REPORTING ABILITY TEST
OQMG 5220

Administrator's Manual

For quick reference:

Name:_____
Test Administrator

Copy number _____

Table of Contents

(1)

INTRODUCTION

This test of reporting ability is being developed as part of a contract between OQMG and the University of Maryland. The final form of the instrument will have wider applicability than picking acceptance testing subjects, although it is on this kind of a group that we are going to validate it first.

Essentially, the task consists of six pictures, each one of which the subject looks at and reports about, under prescribed conditions. The instrument has already been pretested on three different occasions during the past few months. More than three hundred people have served as subjects. Our six pictures are what is left from a group of over three thousand with which we started. We believe that all the bugs are out of the test and you should have no trouble giving it. These pages contain considerable detail to help anticipate as many of your questions as possible. Your further comments and questions will be very much appreciated. Please write them down as they occur to you. If the matter is important, there is a name and phone number which you may call at the back of this booklet. Otherwise, a representative of the University will always be around two or three times a week while these tests are in progress.

APPARATUS

All of the following should be ready before each subject comes into the testing room:

1. A set of six pictures, 8" x 10", mounted on heavy cardboard, and lettered on the back from "A" through "F". (There are three sets of such pictures).

2. A stop watch.

3. A Gray Audograph and microphone.

4. A blank side of a large (half hour per side) Audograph disc.

5. A picture test question booklet.

6. A copy of a rating scale form to which is attached a biographical inventory.

7. A copy of this manual for reading instructions to and keeping record of the subject.

The Stimulus Pictures

The pictures should be treated in the same fashion as any test material which is classified. Probably the single most important factor in the situation is novelty. Neither the test subjects nor anyone not taking the test should see or have access to the pictures at any time, even when the testing period is completed. Only the test administrators are to handle the pictures.

Here is a brief identification for each picture. The letter is printed in ink on the back of the mounting board:

A An ink drawing of a scene in a street (or subway) car.

B A photograph of a couple of sidewalk propogandists and a policeman.

C Photograph of a drawing made outside a ferry landing.

D An ink drawing of a collision in the street.

E Photograph of the interior of a slum tenament.

F Photograph of a drawing of a wedding scene.

SUBJECTS

Our subjects at QMB will consist of the sixty-two officers stationed at Fort Lee who took part in the summerweight uniform study during 1951. We are using these people as a validating group because (1) the questionnaires which they filled out during the course of the uniform study will be of definite value in our later analyses; (2) many of these men know each other; many people at the Board know the men. We will take advantage of these facts in having rating scales filled out by both the subjects and the specialists and experts at the Board. (3) The subjects include a highly specialized expert group: those officers who were or are instructors in QMB school.

(3)

Making appointments for interviews

1. Each of the three test administrators will be responsible for collecting data on about twenty subjects.

2. The test administrator, with the help of a designated officer, if desired, will schedule appointments for one hour periods. Subjects will be contacted as much in advance as possible.

    a. An appointment should be confirmed by telephone on the morning it is to be kept.

    b. A subject who fails to keep an appointment must be contacted again and used. To selectively lose subjects because of appointments not kept would throw a serious bias into the data.

3. The test administrator should allow one half hour between appointments.

    a. Thus, about four subjects can be conveniently seen by one administrator during the day.

    b. No appointments should be scheduled for after 1100. If possible, avoid appointments made after 1600, since the subject may reflect too much fatigue in low motivation.

    c. The entire testing session must be completed at one sitting. Please make this clear to subjects.

4. An appointment log for each administrator may be found on the last pages of this manual.

5. Even though we realize that the test subjects are doing us a considerable service by giving up a portion of a day for the investigation, and even though we would like to make it as convenient for the subjects as possible, one thing must be adhered to: The subject must report to a designated testing room for the test. It cannot be done in his office. The latter would involve too many distractions, and too much work on the part of the administrator.

(4)

PROCEDURE

Summary

The subject looks at each of six pictures for 45 seconds. Immediately after he has seen a picture it is put face down, and he reports as exactly as possible all that he saw in the picture. Right after this verbal report, he is given some True-False questions to answer on the picture he just saw. When this procedure has been repeated for all the pictures, the subject fills out a rating scale and a biographical inventory.

Arrangement of the pictures

1. When the subject walks into the testing room, the pictures should already be arranged. They will be piled face down on the table, in the order in which they will be presented to the subject. Thus, the back of the first picture to be seen will be on the top of the pile.

2. The pictures are lettered on the back from "A" to "F". The order in which they will be presented to the subjects will vary.

TABLE I

Each of the six orders of presentation to be used, together with the identifying series letter.

Series	Top Picture					Bottom Picture
A	A	B	C	D	E	F
B	B	C	D	E	F	A
C	C	D	E	F	A	B
D	D	E	F	A	B	C
E	E	F	A	B	C	D
F	F	A	B	C	D	E

3. In each case you slip the top picture on the bottom to make the next order. The letter that identifies the series is the same as the letter of the top picture, the first to be shown.

4. The test booklets for the subjects have the series letter written in ink in the upper right hand corner. The questions inside the booklet are arranged to conform with the series. It may be helpful to check the arrangement of the pictures with the order of the questions in the booklet before the testing session begins.

5. The letter representing the series identification should be entered in two places:

    a. The space after "Series" in the subject log at the back of this manual.

    b. In the empty block in the upper right hand corner of the biographical inventory. Both of the above entries may be made during or after the testing period with the subject.

6. It is not necessary that you be extremely careful in changing the series for every subject. You can take two or three subjects with the same picture arrangement. The fact that the test booklets are already lettered helps do your job.

Detailed Procedure

The one hour interview with each subject can be broken down into three parts: (A) the verbal report; (B) the written quiz; (C) the ratings and biographical data. (A) and (B) alternate and take up about forty-five minutes, including a short orientation at the beginning. Part (C) will take up about fifteen minutes.

A. Orientation and verbal report

1. Tell the subject what he is to do, then let him read the instructions to subjects on the first page of the question booklet, if he wants to. Include, in your own words, these points:

a. The study is aimed at developing a test which will separate out good reporters from ordinary reporters.

b. Because this is the first time an attempt has been made to develop a test of reporting ability, the scores of the subjects will not be used for any purpose.

c. Except for the test administrator, no one will know which report goes with which subject.

2. Instructions to be read to the subject:

"There are six pictures face down here on the table. When we start, after we've gone through the instructions, pick up the top picture and start looking at it. As soon as I see you pick up the picture, I'll start the stop watch. You will have 45 seconds to look at the picture. At the end of 45 seconds I will say "stop". Put the picture away face down and then tell me as exactly as you can all you saw in that picture.

When you are through making your report there are some True-False questions to answer about the picture. I'll help you with the directions for this part the first time, if you want any points cleared up."

Encourage the subject to ask all questions that might be bothering him. Also ask him to clear up his points of doubt before the actual testing starts.

Using the recorder

1. Use a fresh side for each subject. Make sure he is close enough and talking loudly enough to record properly.

2. Before the subject sees the first picture, ask him to say his name and the appropriate series letter into the recorder. This initial activity somehow relaxes the subject, we have found. After the subject has said his bit, stop the recorder, return it to the beginning, and play back the line or two to make sure the record is working properly.

3. When the subject's 45 seconds exposure to the picture is up, say "stop" and start the recorder immediately. When the subject indicates he has said all of his report, the administrator should say loudly "That's all", or some similar thing, and stop the recorder. This procedure helps the transcriber. Also, by starting and stopping the recorder for each picture, you will get all of the subject's six reports on one side.

B. The written quiz: When the subject has finished making his verbal report on the first picture, and the recorder is stopped, hand him the question booklet, and tell him to read what the possible alternatives mean before turning the page. If he wants some examples, there are a few below that you may use. When he is all set, then tell him he will find all the questions for the first picture on the first page. He is not to look at more than his page of questions at any time.

Two points, we have found, may especially need straightening out:

1. An item is true only when all of it is true. But an item is false when part or all of it is false.

2. NSE and DK are not the same. DK is when the subject doesn't know the answer. NSE is when the picture does not contain the information that would answer the question.

   Example: If the picture shows a man walking with a boy and the statement reads "The boy is walking with his father", then NSE is correct, because there is no way of knowing if the relationship is father and son.

   Example: Picture shows man walking on street with fishing gear. "The man has just returned from fishing" reads the item. Again NSE is correct. Unless there is other evidence (like a string of fish), we don't know if the man is going or returning.

Example: Picture is street scene. There is not any truck in the picture, but the statement says, "The truck in the picture is empty". F (false) is correct answer. The statement is just misleading.

3. The seventeen item quizes shouldn't take longer than three minutes. Tell the subject to hurry if you see him going too slowly. But always give him ample time to finish.

Motivating the subject

1. After the subject has completed the written quiz on the first picture, show him the picture again "to see what you got right and what you got wrong" or "to see how you did". This also helps train the subject.

2. Encourage the subject from time to time, if he did a good job in your opinion.

3. However, do not let him hear his phonograph record, if he wants to, until all six pictures have been recorded. This is important. Doing so slows down the interview tremendously and also seems to change the attitude of the subject unfavorably.

4. There is no need to show him every picture when he is finished with it, unless you feel it is motivating him strongly. At any rate, don't point out his errors or successes if you show it to him.

C. <u>The Rating Scale</u>.

1. This is a vital part of the study. The ratings will be made after all the pictures have been reported on.

2. The rating scale device doesn't look like one. It is a numbered list of every subject. There are two columns ruled out before each subject's name.

(9)

3. Tell the subject, "Here is a list of officers here at Fort Lee (or on the post if you prefer), including yourself. All of these people took part in the summerweight uniform test that the Board (or Quartermaster Board) ran last summer. First, I'd like you to underline your name."

(wait while he does it)

"Now would you please put a check in the column in front of every man's name, whom you know pretty well. These don't have to be your best friends, just good acquaintances."

(wait while he does it)

4. After the subject has checked all the names he is going to check, tell him to go back over the list, and in the column next to the check mark:

   a. make a plus sign for every person checked who did as good or better a job as he did as a subject in last summer's uniform study.

   b. make a minus sign for every person checked who did a poorer job than he did of being an acceptance testing subject last summer.

5. Explanation of the plus and minus ratings.

   a. Emphasize to the rater that the one quality of an acceptance testing subject that we are most interested in is the ability to make good observations of the articles and then report those observations as clearly and helpfully as possible.

   b. The subject may protest that he has no idea how his friends would do as subjects because he has never observed them or talked to them about it. In that case tell the rater to imagine that he was going to pick a number of people for acceptance testing from the names he checked. Let the rater

put the plus and minus signs in front of those names whom he would pick as being the best observers and reporters. If you use this procedure with the subject instead of the preferred one, write "alternate procedure" at the bottom of the rating scale after the subject is finished. Do not use unless you have to.

6. How many names should the subject check?

   a. The number of plus and minus signs that a rater is expected to make depends on how many people he checked.

   b. The man who indicates that he only knows two people, is giving us all the information he can if he makes two signs. The man who knows twenty people is holding back if he only makes two signs.

TABLE II

Rule of thumb table indicating the minimum number of pluses and minuses needed as a function of how many people he knows.

If the rater checks this number	Then he should make at least this number of pluses and minus (half of each).
2-3	2
4-5	4
6-9	6
10-14	8-10
15-19	12-14
20 or over	at least 14

   c. The raters will need a maximum of encouragement and reassurance. But they definitely respond without trouble, and do the job well.

   d. Exceptions to the above rating technique are the QMB officers who took the picture test. We will discuss them as special cases separately.

7. The biographical inventory data form is self-explanatory and should be filled out by the subject as the last thing before he is through in the test. The biographical form is attached to the rating scale sheet. There is a blocked out space in the upper right hand corner. The test administrator should initial this space at the top. Just below his initials he should print the series letter which appears in ink on the front of the subject's question booklet. This may be done after the subject leaves the room.

AFTER THE INTERVIEW

1. The question booklet, rating scale form, and biographical sheet should be paper clipped together and placed in a manila envelope. On the front of the envelope the test administrator should write the names of all the subjects whose data are in the envelope.

2. Before you take the disc off the record machine, say the subject's name, series letter, and "That's all" into the microphone. If both sides have been used, the record may be filed or handed over to designated clerical help for transcription. If only one side has been used, the other side should be used for the next subject.

3. Fill out the rest of the subject log. The log is for keeping track of appointments and identifications. Logs with space for about thirty entries are stapled to these instructions. Please be sure that the logs and this manual are returned at the end of this study.

4. The designation "code no." on the first line of the log refers to the number which the subject has on the rating scale form.

5. Rearrange the pictures in order for the next subject before replacing them in their envelope. The letter on the back of the top picture identifies the series.

(12)

How to get additional help in answering your questions:

1. A University representative will be at QMB two or three times a week during the course of this study.

2. If you have a question which is important, call the University of Maryland at College Park:

   WArfield 3800
   Extension 403

   Ask for Mr. Abelson. If he is not there, either Mr. Hembree or Mr. Parker at the same extension should be able to help you.

Initials ____________
Test admin.

Picture Test
SUBJECT LOG

1. Name ______________________ Rank _______ Code No. _______

2. Post telephone number _________

3. Appointment made for: Day ________ Date _______ Hour ____
   (1) Appointment confirmed on: Date _______
   (2) If not kept at above time, when was it kept?
   Day ________ Date __________ Hour ____

4. Room used A__ B__. Picture series___. Record side 1_2_

. . . .

1. Name ______________________ Rank _________ Code No. _______

2. Post telephone number ________

3. Appointment made for: Day ________ Date _______ Hour ____
   (1) Appointment confirmed on: Date _______
   (2) If not kept at above time, when was it kept?
   Day ________ Date __________ Hour ____

4. Room used A__ B__. Picture series___. Record side 1_2_

. . . .

1. Name ______________________ Rank _______ Code No. _______

2. Post telephone number _________

3. Appointment made for: Day ________ Date _______ Hour ____
   (1) Appointment confirmed on: Date _______
   (2) If not kept at above time, when was it kept?
   Day ________ Date __________ Hour ____

4. Room used A__ B__. Picture series___. Record side 1_2_

. . . .

1. Name ______________________ Rank _______ Code No. _______

2. Post telephone number _________

3. Appointment made for: Day ________ Date _______ Hour ____
   (1) Appointment confirmed on: Date _______
   (2) If not kept at above time, when was it kept?
   Day ________ Date __________ Hour ____

4. Room used A__ B__. Picture series___. Record side 1_2_

PROCEDURAL CHANGES

These changes take precedence over any conflicting material in the test administrator's manual.

Subject Identification

1. Ratings will be anonymous to help insure freedom of subject when rating.
2. As part of the introduction to the interview, the subject will be asked to pick a six digit number.
3. That number will be written on:
    a. Subject log
    b. Question booklet
    c. Rating scale
    d. Biographical form
4. The number will be used also to identify the subject's Audograph record.
5. After the subject has left the room, the administrator will underline his name on the rating scale.
6. A letter signed by Colonel Jordan, assuring the subject that his report will be kept confidential, will be shown each subject at the beginning of the interview. No one will have access to the report except the administrator and University of Maryland personnel.

Administrator Rating of Subject

1. After subject has left the room, the administrator will use the unused portion of the record for comments.
2. In lower right hand corner of the rating scale sheet, the administrator will write a plus sign if the subject did a conscientious job of rating, a zero if the subject was average, a minus sign if the subject was poor.

APPENDIX C

Scoring Materials

Instructions for Scoring Reports

Notes for Scoring Each Picture not Specifically Covered in General Instructions

A Sample Scored Report for Each Picture

Question Booklet Used in Recognition Form of the Test

Scoring Key for Recognition Test

INSTRUCTIONS FOR SCORING REPORTS

General objective: To give credit proportional to the specificity of the report. To note the items in the report which are not given in the picture (inferences, ambiguities).

Aids to scoring:

1. Have the appropriate picture in front of you.
2. A set of notes have been included for each picture. These indicate the procedure for scoring special items not covered by the general description of each index.

    a. Read the notes over once before you begin scoring.
    b. Use the notes for reference during scoring.

Each report is scored for four different indices.

1. Detail (D)
2. Clothing detail (C)
3. Unreasonable inferences (I)
4. Ambiguities or perceptual shifts (A)

Mechanics:

1. Underline once each word or phrase that scores for detail.
2. Underline twice each word or phrase that scores for clothing detail.
3. Parentheses around each word, phrase, or sentence which is one unreasonable inference.
4. Double parentheses around each word, phrase, or sentence that is either an ambiguity or false detail. (False detail is detail added to the picture which was not originally in it.)

Scoring Detail

The more specific the report, the higher the score.

1. Any object named correctly is scored for detail.
2. Any word or phrase which modifies the object to add specificity is also scored for detail.

3. The location of an item either relative to something else in the picture or relative to the person viewing the picture is scored for detail.
4. Some words describe situations rather than objects but are included in the detail category: wedding, accident, discussion.

Exceptions to the above:

1. A repeated object scores only the first time. If it is modified the second time, the modifier alone is scored.
2. If the reporter names something incorrectly, and immediately corrects himself, the first incorrect naming is disregarded.
3. Some items are not scored at all. The notes on scoring each picture indicate which ones.

Scoring Clothing Detail

1. Any item of clothing is included in this group.
2. Certain items such as eyeglasses, hair ribbons, wrist watches, etc., are also scored for clothing. See notes for each picture.

Scoring (Unreasonable) Inferences

1. Reasonable inferences are scored the same way as detail.
2. Unreasonable inferences are scored separately.
3. The word inference alone will connote unreasonable inference.

Examples:

1. Calling a man a policeman because he is wearing a policeman's uniform is a reasonable inference.
2. In the wedding picture, identifying bride, groom, best man, etc., is reasonable inference.
3. "The truck driver is looking out the window." "Looking" might be considered reasonable and scored for detail. But "The truck driver is looking at the lady". "Looking" is probably unreasonable in this sense. See notes for pictures.
4. Anything described as typical is unreasonable.
5. Usually, comments on the weather, a person's nationality, religion, destination, speed of traveling are unreasonable inferences.

Scoring Ambiguities

1. These are usually errors of commission.
2. False detail is also included in this category.

Examples:

1. Misspelling a name or a sign.
2. Locating people or objects in wrong relative position.
3. Changing the appearance of an object (motorcycle becomes bicycle).
4. Adding something to the picture that is not in it.

NOTES ON PICTURE A

Key: D-Detail, C-Clothing, A-Ambiguity, I-Inference

General Setting, Location, Outside Window, etc.

1. Do not score any of the following pertaining to the location of the street car: "at a station", "resting", "stopped", "passing".

2. "The subway platform is outside the window." Reasonable I. Score by underlining as noted.

3. "Picture was taken at Dykeman Street." Underline as noted. If the word 'photograph' is used, score the phrase as A.

4. Referring to picture itself: Score only one of these if more than one appears: "this sketch", "drawing", "cartoon", "ink drawing". No score for the word "picture". Notice that ink drawing is scored as two separate words.

5. If location is given as "New York", do not score. If "Brooklyn" or "Bronx", score I.

6. Any one of these is scored as noted. Where more than one appears, only one is scored: elevated, elevated train, trolley car, interurban car, street car, train, express train (scores twice). Underground railroad, no score; Northeast express, score A.

7. For name of car "A N Courtland Park", "Van Courtland Park", "One assumes that this is the 7th Ave. express destined for Van Courtland Park", no score.

8. "inside" or "interior of" a street car, etc. No score for inside or interior.

9. Street car is "stopped at Courtland Park". Score both I and A.

10. "Lots of trees outside" or "foliage in full". Considered reasonable inference, and thus is scored as noted.

11. The time is "twenty to four". Scored as a unit. In the "afternoon", no score. In the "morning", score I.

12. "typical" as a modifier of anything scores both it and its noun as I.

13. "No papers littered the floor." Score as noted. Usually negative statements are not scored.

14. "There are no cross seats on the car."

15. "Subway is above ground", no score.

The Advertisements

1. "There are four ads up on the ceiling over the people's heads." Score as noted. No score for locating the ads, except relative to each other.

2. "sign", "poster", "advertisement". Score any one once. If two different words appear, score only one.

3. "three or four ads" score D for ads, but nothing for number because two numbers named.

4. "advertisements for various products" no score for "various products". Redundant.

5. "lucky rakes", "Lucky Strike" A.

6. "Smoke Lucky Rakes cigarettes"

7. "Gosling soap", "soap", "picture of a swan" score I, "picture of duck" score I.

8. "Idlewild Camp" or "Camp Idlewild"

9. "Vote for McGinnis for Alderman", "voting". The word "alterman" (t instead of a d) is scored as though it were correct.

The People

1. The word "people" is not scored if "men" and/or "women" is included in the report. If not, score people.

2. When "men" or "women" is scored early in the report, and the number of them added later, score only the number the second time, not the word "men" again.

3. "Two men standing."

4. "Two women, the rest are men", score as noted

5. "People ordinarily dressed" or "dressed for weekday", score either one (not both together) as noted for both D and C (red pencil).

6. "people dressed for shopping tour" score I

7. "there are seven people" score A (baby)
"there are seven adults"

8. There are "six or seven people" no score

9. Of the two men who are standing:

   a. "having a discussion" or "talking" score as noted. But "arguing" scores I as does any synonym of it: "heated discussion" score I.

   b. "One of the men standing appears to be quite dark"

   c. "razor in his belt": score A, "razor in his hand"

   d. Describing the white man as a "workman", "laborer" scores I, wearing "Work clothes", calling him a "mechanic" scores I.

   e. "badge" on man standing up. "Sleeves rolled up", scores for both D and C.

10. Man reading newspaper

   a. "One of the men is reading the Boston Globe."

   b. "a Boston newspaper" "newspaper"

11. The women

   a. Speculation on age of either woman is I. Calling elderly or younger, etc., is also I.

   b. eyeglasses, purse, hair ribbon, score for D and C

   c. "one of the women is holding a child" no extra credit for holding or carrying.

12. Bearded Man

   a. "Elderly man with beard". If beard is mentioned, don't score "man". Elderly scores D.

   b. Beard does not score for C

c. "One of the men needed a shave" score both I and A.

d. "Jewish Rabbi", "Jewish", "Rabbi", all score I.

e. Wearing a "derby hat" okay for both D and C.

13. Man on extreme right

a. "Fat", "Obese"

b. "Chinese" scores I

c. "bald headed"

d. "appears to be sleeping" score I, also "dozing" is I

e. sleeping "with his head on his shoulders" no score.

NOTES ON PICTURE B

Key: D-Detail, C-Clothing, A-Ambiguity, I-Inference

General Setting

1. If subject sets scene in front of White House, no score for White House. All subjects not equally familiar with locale.

2. "In front of White House" okay to score D for relative position.

3. "park" or "estate" "park or estate"

4. This is a "photograph" "(Glossy) Photograph" glossy scores A

5. "Sidewalk scene" "Wide sidewalk" If wide appears later on alone, don't score sidewalk twice.

6. "High iron picket fence" "Steel" equally okay

Weather and Season

1. "It has been raining" score any variant the same.
"It is raining" scores I.
"It does not appear to be raining"

2. When report says "it is raining" and offers damp sidewalk as proof, no score for damp sidewalk. Score "Damp sidewalk" only if offered as a fact by itself or in conjunction with "It has been raining".

3. "Sun appears to be shining"

4. If spring or summer is mentioned as time of year, score for D IF a reason is given: Trees in bloom, or people not wearing heavy coats.

The Pickets, Policeman and Tax

1. "Two men, one talking to a policeman"
2. Talking or discussion okay for man and policeman, but arguing, fighting, etc., is scored I.
3. "protesting", "advocating repeal", "picketing", "demonstrating"
4. "Two men advocating repeal of the 20% amusement tax"
5. Score "predominant object" when it refers to postcard, or pickets.
6. "papers", "pamphlets", "literature" all okay for D when referring to what the picket is holding in his left arm.
7. Neither picket is leaning on or holding up the post-card. This is scored A.
8. "Holding a placard" or "carrying" okay for man with sign on stick.
9. "One man is holding replica of a postcard upraised" score sentence A.
10. Picket with hat on has "package in his hand" score A.
11. The reference that the subject makes to the signs must be an exact quotation, or the reference scores A, even if the meaning is clear.
    Score signs this way if quoted correctly:
    "We demand repeal of the 20% amusement tax."
    "The war is over but the 20% amusement tax lingers on."
12. "The postcard was addressed to Congress, U S A"
    Do not score postcard again, of course, if it has been mentioned once.
13. No score for naming size of postcard, but if size is obviously too big: 4 by 6 feet or larger, score A.

Dress and Other People

1. "Two more persons, a man and a woman" no score for persons if sexes are named.
2. Woman is looking through fence okay for D. Reasonable

3. "No hat" or "bareheaded" score C and D

4. "Holding an umbrella" No score for C on umbrella, just D.

5. Policeman's shoes and/or raincoat black okay for C and D. "light colored" or dark colored okay if correct for C and D.
Otherwise naming colors for anybody's clothing is I.
For the two men in back, score I if subject says single or double breasted suits.
For man carrying sign on stick: "suit coat" okay but "suit" is I.

NOTES ON PICTURE C

Key: D-Detail, C-Clothing, A-Ambiguity, I-Inference

General Setting

1. Season of year: "spring" or "summer" okay for D
2. There are "two convertibles" "ten cars" waiting to be loaded. When report says "9 or 10 cars" or "2 or 3 trucks" etc., no score, because of OR.
3. Do not score "Matapeake" or "Gade" or any of their variants. Some error may have been introduced by the fidelity of the transcribing machine.
4. "Ferry on right would dock soon" It is assumed that the word ferry has been scored once.
5. "Concrete boulevard" or divider etc., okay for D
6. "Good weather" or "sunny day" okay because of shadows and scattered clouds.
7. "Ferry Pier" "Ferry wharf," ...dock, landing, approach, slip. Score any one once.
8. "One ferry has just left" score A JUST is No good
   "One ferry has just docked" score A
   "One ferry is loading" "One is coming in"

People

1. Calling the woman or girl the wife or girl friend of any of the men is I.
2. Identifying which man belongs in which car, or who is the driver of the convertible is I.
3. "One of the men is bald" score A "partly bald" or "Balding" score D.
4. Saying that the convertible is going to or coming from vacation, or both, is score I.
5. "They were going to go on the ferry boat"

6. "One man at least with no hat" both score C and D "None of the men wore hats" (extra detail) score all three C and D.

7. Any total number of people named as being in the picture is not scored, whether right or wrong.

8. Clothes: "White shirts" Otherwise no colors may be reported. Any colors score I. However, light or dark, if correct, is underlined for C and D.

9. Eyeglasses score C and D. D only for golf bags, fish net, hat box

10. There are three suitcases, one hat box. Four suitcases are also acceptable.

11. "Crabbing net" Basketball net or fly net are I.

12. "several suitcases" "Portable typewriter" scores I.

The Situation

1. "Convertible" or "Convertible coupe" or "Roadster"

2. "The cars have locked bumpers" "Changing a tire" is A

3. "apparently trying to unlock bumpers"

4. "Truck driver is watching the activity at the convertible" is I.

5. "Top down"

Miscellany

1. "Man on ferry is directing cars"

2. "toll gate at a ferry"

3. "Ferry is going to the Eastern shore of Maryland." No score.

NOTES ON PICTURE D

Key: D-Detail, C-Clothing, A-Ambiguity, I-Inference

General Setting

1. The time is "five minutes after ten". 10:03 or 10:07 also okay, but 10:10 is A. When subject gives two possible times, no score. Morning or evening are I.
2. "New York State license", "1919", "April 14" New York City is I.
3. "It is or has been raining." Clear day is A "Puddles"
4. "Corner of second avenue." Second street is A
5. "Department store" is I
6. "Sign reads owl cigars" is A. "Sign for owl cigars"
7. "Shoe company is conducting a sale" is A
8. "Street on left is on a hill"
9. "Concrete office building" (or any substance) is I
10. "You can't read what the other street is."

The Accident

1. Any statement of how the accident happened, "car slid down a hill" or skidded, etc., is I
2. "Accident happened a short time previously." But "accident just happened" or "happened at five after ten" is I.
3. "Tracks where the auto has skidded" is I.
4. Brand name of car is I. "Car not damaged much", or "damaged very much" either one is I.
5. "Touring car" no score for touring. Roadster, truck, ambulance are all A. Sedan also A.

6. "Injured man was driver." Passenger or pedestrian is A

7. "This picture is about an accident"

8. "Wreck" is I.

9. "Street car hit car" Score only first time.

10. Car was hit on left. okay "Front" score A. Side okay.

The People

1. "newsboy" or "paper boy"

2. No score for boy is waiting to cross street, or is crossing street. Score I for boy running or boy going to the accident.

3. "Five people" lady across the street is A all male

4. "No passengers on street car"

5. "Wheeler insurance company"

6. What the policeman is doing to the man is I. "giving first aid" "identifying him"

7. Policeman's badge is 23 or 25. Either one okay for D.

8. "Man injured." But dead or seriously injured or unconscious are all I.

9. "Motorman" or "conductor" (score one once)

10. "Man was thrown from his car" is I.

11. "Holding victim's hat" hat scores C and D. Victim not scored twice if mentioned already.

12. "Sergeant kneeling." Bending is A also standing over. Two policemen is A.

NOTES ON PICTURE E

Key: D-Detail, C-Clothing, A-Ambiguity, I-Inference

1. "One room flat" apartment okay, bedroom, no score. Cold water apartment is I.
2. Don't score object on other side of window.
3. Don't score whatever is being ironed.
4. There is no comb on the dresser. If subject puts it in, score A.
5. white spread. White is I. "chenille bedspread", "candlewick bedspread"
6. "colored woman" "ironing" clothes
7. "Part of a bed"
8. "coat hangar" lying "on the bed"
9. "man's hat" score hat for C and D; man for D
10. Dresser, vanity, bureau, credit for any one.
11. One of the "glasses has water". Water okay for D.
12. "wrist watch" "bracelet" D only no C
13. "Clothes on bed dampened for ironing." score I
14. "Artificial flowers"
15. "One ironed article on the bed"
16. "Room serves as bedroom and kitchen" score I.
17. No score for what has been ironed.
18. Age of woman is I. "heavy or stocky" okay for D.

NOTES ON PICTURE F

Key: D-Detail, C-Clothing, A-Ambiguity, I-Inference

General Setting

1. "small town", "large town", "residential area", any one of these scores I.
2. "this is a picture" or "this is a photo" no score.
3. No score for naming artist, but score D if name is located in right place (lower left of picture).
4. "summer" or "Spring" scores D only if some other fact is mentioned with it. Otherwise no score.
   "It is summer because of the foliage on the trees"
   "It is summer because no one wears a hat"
5. "June wedding" scores I.
6. "Sunny day" (okay because of shadows)
7. "Some trees to right of church" score A
   "A tree to the right of church" (church is only underlined if it has not been underlined before).
8. "Sky is overcast" or "It will rain later" score I.
9. "Small white church"
10. "Scene is right after a wedding"
11. "Concrete steps" "cement steps"
12. "protestant denomination"
13. "recent make auto" "postwar auto" but "new auto" scores I. Any specific brand is I.
14. "Car is predominant object in picture" score I. Saying who owns the car or that the car will be used for the honeymoon is scored I.
15. "picture taken on steps of church"
16. There are 17 people. 6 women, 7 men, 4 children, 2 boys, 2 girls.

Bridal Party

1. One each for the first mention of bride, groom, best man, maid of honor or bridesmaid, minister or preacher.
2. Relation between members of bridal party and any of the spectators is scored I.
3. Relations within the bridal party scored I.
4. "several families represented" score I.
5. Flowers and eyeglasses score both C and D.
6. "best man congratulating groom" scores I
   "best man shaking hands with groom" "holding hands"
7. Bridesmaid has "large hat" C and D "floppy hat"
8. Bridesmaid "good looking" or "pretty"
9. Minister wears "dark suit" C and D
10. Best man "Light suit" score A
11. Minister "holding a Bible"

Photographer

1. "Camera" (okay even if not in picture) "tripod"
2. "Photographer is taking a picture" score first mention of either photographer or taking a picture. Not both.
3. Photographer "bent over" "stooped over"
4. "Holding safety plate" also okay if he is holding exposure plate, a book, etc.

Dog

1. "cocker spaniel". Any other breed scores A.
2. Who dog belongs to is I.
3. "Dog on the steps"

4. "Black dog" scores A. If color mentioned, white spot must also be mentioned.
"Small black cocker spaniel with white spot on chest"

5. "Dog is predominant" or "Dog is center of attraction"

6. "puppy"

7. "several people are trying to coax the dog". Score once the first mention of coax or any reasonable synonym. In example above, several people gets no score.

8. The reason why the dog is being coaxed or is the center of attention is not scored. Examples: "to get it out of the picture" or "to make a photograph of the bridal group"

9. "small boy motioning with his finger to the dog"

10. "A man appears to be trying to get the dog out of the way"

11. The man with the hat is "shaking it at the dog" scores I.

Spectators

1. "two girls, and two boys"

2. "all the men have their ties and coats on" men scores only if not mentioned before. "All" becomes an important word in this context.

A SAMPLE SCORED REPORT FOR EACH PICTURE

Picture A

Serial No. 333196 Series E

This picture was a scene on the inside of a streetcar ....it's not a photograph but a drawing of a picture.... it was an avenue express car going to ((Cortland Street)) I believe -- as much as I could see....it still had the old advertising posters in the streetcar, advertising cigarettes, soap, and ((hair tonic)) and a political advertisement....there were two women on the streetcar -- one holding a baby....and five men....one of the (men was sleeping)....((three of the men were standing))...there was a drugstore on the corner looking out of the streetcar ....there was a clock that said -- had a sign that said Dykeman street....the clock, as I recall, it was about twenty minutes of four....all of them (appear to be a poor class of people) -- (laborers or for the most part).... one man was colored....that is all.

Picture B

Serial No. 159447 Series D

there's an iron fence ah around the park...a side-walk scene along side of a park which is bounded by a iron or uh metal fence....there were five men excuse me there were five men in the picture one a policeman the other four civilians one woman ah a woman and a man rather ah walking together...it must be raining or must have rained and is just drizzling cause the man is holding the umbrella (for the woman)...the time of year it would be late spring or early summer the trees are all in full foliage...(the picture is comparatively recent)...the cars are parked across the parkway are of (recent make)...there are two signs in the picture describing or talking about the 20% amusement tax...one sign is a blown up version of a ((one cent postage stamp)) addressed to Congress U S A... this regards to the amusement tax....the other sign is ((on the park fence)) and it states that ((the war is over but that they still have the 20% amusement tax))...the man talking to the policeman is plainly dressed he has a sport coat and slacks on...(the other three men are in suits) ...the policeman has a raincoat on....the time of day would be (early afternoon)....

Picture C

Serial No. 133059 Series C

this scene seems to be a toll gate at a ferry.... Meadowpeak ferry...the cars are lined up....the policeman is motioning them forward...the stop gate is up however one of the cars with a great deal of baggage and fishing net in back is stopped ((apparently for a flat tire))...there are nine cars in front of the car stopped...there is another ferry approaching the harbor...there is a ((moving)) truck beside the stopped car....it's a convertible car top down ...that's all.

Picture D

Serial No. 135974 Series A

it's a scene of an accident...uh street car number 186 has ran into a New York vehicle that has a license number 1919...(evidentally the year of the car as it is rather dilapidated)...there's a man lying on the ground and two persons by him a ah policeman with badge number 23 and the other is apparently a motorman...a there's a little newsboy on the left hand corner with ah (standing by ah running rather by an by uh) a lamp post or clock post there's a clock on top of it rather than a lamp...across the street is a building rather large...the ((left corner)) of the building it's the telephone office....dentist office is on the second floor....there's a sign rather large banner stretched across the building it says sale starts today April the 14th...on the right hand corner side of the building there's a sample shoe store ...((on top of the building there's a big sign that ah says)) Owl...that's all for picture number 4.

Picture E

Serial No. 331225 Series D

the picture is a (hotel room) with for (two people)... there is a colored woman ironing what appears to be sheets though it could easily be a shirt...ah bed had three chairs two tables visible and the mirror reflected another table or a ah dressing table which apparently is a woman's dressing table because the ((one in the picture appears to be that of a man))...on the bed there is a coat hangar ah a wire coat hangar and ah handkerchiefs and another large colored piece of cloth could very easily have been a table cloth...there's a sink in the kitchen its in the room rather that ah is like a kitchen sink...there's a picture of a woman on what appears to be the table a vase and a flowers...in the mirror is reflected a man's shirt hanging and a sleeve of what appears to be a jacket a dark colored cloth...

Picture F

Serial No. 987631 Series A

this scene is a church scene...there's a wedding (coming out of the church)...standing on the top steps of the church are ((three men and a woman))...down below them is a little ((black dog))which several people are trying to call away... a photographer is trying to take a picture of the wedding party...the minister has a book in his hand...to the right are several spectators including ((one small boy)) and two small girls...there are also several men and women in this location...there is a (new automobile) parked in front of the church...there are ((some trees))to the right of the church...that is all.

Series A

PICTURE TEST
Question Booklet

As you know, you are helping us develop a test which can be used to screen acceptance testing subjects. Your help is very much appreciated and you should be reassured of two things: (1) At no time will anyone but the test administrator know which test report goes with which person. Your name will be removed and a code number substituted. (2) The score you get on this test will not be entered with your name anyplace, nor will it make any difference in your being asked to be an acceptance testing subject in the future.

The first part of the test consists in telling as exactly as possible all that you see in each of six pictures. You make your report right after each picture has been exposed to you. The second part, the quiz in this booklet, comes right after each verbal report. It is designed to let the keen observer make up for not being able to talk out all he has seen.

On the following pages you will find a set of 10 to 17 statements for each picture. To the left of each statement are four abbreviations: NSE T F DK. Circle the abbreviation that best answers the statement. Here is what the abbreviations mean:

NSE Not Sufficient Evidence. It is not possible to tell from the picture if the statement is true or false.

T The statement is entirely True.

F The statement is partly or entirely False.

DK Don't Know. You don't remember this part of the picture well enough to give an answer.

Notice that DK and NSE are quite different. DK means that YOU don't remember what was in the picture. NSE means that the PICTURE did not contain a sure way of deciding if the statement is True or False.

You may refer back to the above key at any time. If you have any questions they can be answered by the test administrator. Please ask all questions before you begin.

IT IS VERY IMPORTANT THAT YOU DO NOT LOOK AHEAD IN THIS BOOKLET AT ANY TIME. If you see the statements before you are supposed to, your score would be partly invalidated.

PICTURE A

NSE	T	F	DK	1.	The man reading the newspaper is bald.
NSE	T	F	DK	2.	The man who is sleeping is Chinese.
NSE	T	F	DK	3.	One of the women wears glasses.
NSE	T	F	DK	4.	The Lucky Strike ad is on the far left.
NSE	T	F	DK	5.	The woman with the infant on her lap is its mother.
NSE	T	F	DK	6.	The Negro has a razor in his hand.
NSE	T	F	DK	7.	The woman who wears glasses has a shopping bag on one arm.
NSE	T	F	DK	8.	The White man who is standing has the same pattern on his necktie as the Negro does.
NSE	T	F	DK	9.	The man with the beard is wearing a bowler hat.
NSE	T	F	DK	10.	The ad for Swan Soap says "cheaper, better, purer".
NSE	T	F	DK	11.	The standing White man is more plainly dressed than the Negro.
NSE	T	F	DK	12.	The woman holding the baby has a ribbon in her hair.
NSE	T	F	DK	13.	The Negro is wearing shoes with no laces.
NSE	T	F	DK	14.	This car is part of an express.
NSE	T	F	DK	15.	It's a sunny day out.
NSE	T	F	DK	16.	There is an empty seat between the two women.
NSE	T	F	DK	17.	There are some leather straps hanging down for standing passengers.

PLEASE DO NOT LOOK AHEAD

PICTURE B

NSE	T	F	DK	1.	There are five men in this picture.
NSE	T	F	DK	2.	Four of them are wearing hats.
NSE	T	F	DK	3.	The model postcard says on it at the top: "The war is over but the 20% amusement tax lingers on."
NSE	T	F	DK	4.	Two of the men in the picture are wearing glasses.
NSE	T	F	DK	5.	The couple in the background are carrying a closed umbrella.
NSE	T	F	DK	6.	The only man wearing a raincoat is the policeman.
NSE	T	F	DK	7.	The man with the papers in his hand is wearing a sport shirt open at the collar.
NSE	T	F	DK	8.	The trees are bare; they have not yet begun to blossom.
NSE	T	F	DK	9.	All the people in the picture are looking at the three men in the foreground.
NSE	T	F	DK	10.	The placard says on it, "We demand repeal of the 20% amusement tax".

PLEASE DO NOT LOOK AHEAD

PICTURE C

NSE T F DK 1. There are no prewar cars in the picture.

NSE T F DK 2. There are four people in the immediate foreground.

NSE T F DK 3. Two men are bending over the left front of the convertible in the foreground.

NSE T F DK 4. The two men working on the front of the convertible have their sleeves rolled up.

NSE T F DK 5. One of the men working on the convertible is its driver.

NSE T F DK 6. The woman is standing on the left side of the convertible.

NSE T F DK 7. The woman in the convertible has her sleeves pushed up.

NSE T F DK 8. A man's suit jacket is lying over the luggage in the back of the convertible.

NSE T F DK 9. One of the items in the back of the convertible is a golf bag.

NSE T F DK 10. One of the truck drivers in the right lane is looking at the woman standing in the car.

NSE T F DK 11. The ferry that is loading is at the left landing.

NSE T F DK 12. The name of the ferry is the Matapeake.

NSE T F DK 13. The policemen in the picture are directing the left lane of traffic.

PLEASE DO NOT LOOK AHEAD

PICTURE D

NSE T F DK 1. There are puddles of water in the street.

NSE T F DK 2. The street car is owned by United Trucking Lines.

NSE T F DK 3. The street leading into the intersection (West to East), is a hill.

NSE T F DK 4. It would be safe to say, from evidence in this picture, that the locale is New York City.

NSE T F DK 5. The street car has a small advertisement on the front.

NSE T F DK 6. The date when this accident happened is April 14.

NSE T F DK 7. The vehicle that is in the accident with the street car is a small truck.

The damage to the vehicle in the accident with the street car consisted of:

NSE T F DK 8. A bumper which has partly fallen off.

NSE T F DK 9. A damaged radiator grille.

****

NSE T F DK 10. The damage to the street car includes a broken trolley pole.

NSE T F DK 11. One of the men near the victim is a policeman.

NSE T F DK 12. One of the men near the victim is the street car motorman.

NSE T F DK 13. One of the men near the victim is a movie usher.

NSE T F DK 14. The outside of an accountant's office is visible in the picture.

NSE T F DK 15. There is a bowling alley sign in the picture.

NSE T F DK 16. The sidecar on the policeman's motorcycle has a white stripe painted on it.

NSE T F DK 17. The street car looks full of passengers.

PLEASE DO NOT LOOK AHEAD

PICTURE E

NSE	T	F	DK	1.	This is a scene in a one room flat.
NSE	T	F	DK	2.	The Negro woman is doing ironing for someone else.
NSE	T	F	DK	3.	She is using the kind of iron that has to be heated on a fire.
NSE	T	F	DK	4.	The bed is made.
NSE	T	F	DK	5.	There is a spread on the bed.
NSE	T	F	DK	6.	There are three chairs in the picture.
NSE	T	F	DK	7.	They are all kitchen type chairs.
NSE	T	F	DK	8.	There is a soap dish but no soap in it.
NSE	T	F	DK	9.	There are two iron clothes hangars on the bed.
NSE	T	F	DK	10.	A man's hat is in the picture.
NSE	T	F	DK	11.	A man's shirt is in the picture.
NSE	T	F	DK	12.	There is a vase with flowers on the table.
NSE	T	F	DK	13.	The table has a drawer in it.
NSE	T	F	DK	14.	There are some articles of food on the table.
NSE	T	F	DK	15.	There is a man somewhere in the room, but he is not in the picture.

PLEASE DO NOT LOOK AHEAD

PICTURE F

NSE	T	F	DK	1.	All the children are on the right side of the picture.
NSE	T	F	DK	2.	The groom is wearing a double-breasted suit.
NSE	T	F	DK	3.	The photographer is holding a mechanical birdie in his left hand.
NSE	T	F	DK	4.	The bride has a veil over her face.
NSE	T	F	DK	5.	The clergyman is wearing a turned around collar.
NSE	T	F	DK	6.	The dog in the picture is all black.
NSE	T	F	DK	7.	Both of the church doors are open.
NSE	T	F	DK	8.	The clergyman is wearing a double-breasted suit.
NSE	T	F	DK	9.	The groom has a flower in his buttonhole.
NSE	T	F	DK	10.	Nobody else in the picture except the groom and the best man have flowers in their buttonholes.
NSE	T	F	DK	11.	None of the men in the picture is wearing a hat.
NSE	T	F	DK	12.	The dog is owned by the boy who is beckoning to it.
NSE	T	F	DK	13.	The maid of honor who is standing next to the bride is the only other woman with a long gown.
NSE	T	F	DK	14.	One of the boys is wearing a striped polo shirt.
NSE	T	F	DK	15.	There is one woman with gray hair and one man with gray hair in the picture.
NSE	T	F	DK	16.	The photographer's camera is mounted on a tripod.
NSE	T	F	DK	17.	There is a small girl with a bag in her hand.

PLEASE DO NOT LOOK AHEAD

SCORING KEY FOR RECOGNITION TEST

Items may be scored true (T), false (F), or not sufficient evidence (NSE). The items are then divided in categories representing the proposed indices of reporting ability: detail (D), clothing detail (C), ambiguity (A), or inference (I). Of the four indices, three are mutually exclusive. Clothing detail alone is scored with one of the other indices, never by itself.

PICTURE A

Item Number	Key	Index
1	NSE	I
2	NSE	I
3	T	D
4	F	A
5	NSE	I
6	F	A
7	T	D
8	F	CA
9	T	CD
10	F	A
11	T	CA
12	T	CD
13	T	CD
14	T	D
15	NSE	I
16	T	D
17	F	D

PICTURE B

Item Number	Key	Index
1	T	D
2	T	CD
3	F	A
4	F	D
5	F	A
6	T	CD
7	F	CD
8	F	D
9	F	A
10	F	A

PICTURE C

Item Number	Key	Index
1	NSE	I
2	T	D
3	T	D
4	T	CD
5	N	I
6	F	A
7	T	CD
8	F	CA
9	T	D
10	NSE	I
11	T	D
12	T	D
13	F	A

PICTURE D

Item Number	Key	Index
1	T	D
2	F	A
3	T	D
4	F	I
5	F	D
6	T	D
7	F	D
8	T	D
9	F	D
10	F	D
11	T	D
12	T	D
13	F	D
14	F	D
15	F	D
16	F	D
17	F	D

PICTURE E

Item Number	Key	Index
1	NSE	I
2	NSE	I
3	F	A
4	T	D
5	T	D
6	T	D
7	F	A
8	F	A
9	F	A
10	T	CD
11	T	CD
12	T	D
13	T	D
14	F	D
15	N	I

PICTURE F

Item Number	Key	Index
1	T	D
2	T	CD
3	F	A
4	F	CA
5	F	CA
6	F	A
7	T	D
8	T	CD
9	T	CD
10	T	CD
11	T	CD
12	NSE	I
13	T	CD
14	T	CD
15	T	D
16	T	D
17	T	D

APPENDIX D

Biographical Data Sheet and Summary of Biographical Data Used in Analysis

BIOGRAPHICAL DATA

1. Name_____ 2. Rank_____

3. Total time in service (active duty) _____

4. Public school education: circle highest grade completed

   1 2 3 4 5 6 7 8 9 10 11 12

   College education. Circle last year completed

   0 1 2 3 4 5 6 7

   College degree if earned_____

   Field of specialization in college _____

5. MOS: number _____ Title _____

6. If you have completed QMB school, write date of completion here: _____. Which course?_____

   _____.

7. If you are now a student at QMB school, state expected date of completion: _____.

8. If you have been or are an instructor at QMB school, when did that assignment start: _____.

   If you are no longer an instructor, note date of termination of assignment: _____.

9. If your present assignment is not covered by #6, 7, or 8 above, please state it here:

   _____

   _____

10. If last summer's uniform test was not the only acceptance study you have taken part in at Fort Lee, please list below or on the back of this sheet (1) the names of the other items that you have tested or are testing; (2) the approximate dates of each of the studies.

SUMMARY OF BIOGRAPHICAL DATA USED IN ANALYSES

Subject Number	Rank	Years in Service	*Education	Instructor in Quartermaster School
1	Major	14	2C	yes
2	Major	10	D	yes
3	Major	11	HS	no
4	2nd/Lt.	3	D	no
5	1st/Lt.	9	D	no
6	2nd/Lt.	3	D	no
7	1st/Lt.	4	D	yes
8	Lt/Col.	12	D	yes
9	1st/Lt.	4	D	no
10	1st/Lt.	9	HS	yes
11	Major	12	HS	yes
12	WOJG	12	HS	yes
13	Major	11	D	yes
14	Lt/Col.	11	1C	no
15	Major	11	D	no
16	1st/Lt.	11	HS	no
17	Capt.	13	D	yes
18	Major	17	HS	yes
19	Capt.	9	3C	yes
20	Capt.	6	D	no
21	Lt/Col.	10	D	yes
22	Lt/Col.	12	2C	yes
23	1st/Lt.	6	D	yes
24	Capt.	(no other information supplied)		
25	Lt/Col.	11	2C	yes
26	Capt.	5	5C	yes
27	Major	12	3C	yes
28	Capt.	9	3C	yes
29	Capt.	11	1C	yes
30	Major	10	2C	yes
31	Capt.	9	D	yes
32	1st/Lt.	10	D	yes
33	1st/Lt.	13	1C	no
34	Capt.	11	HS	no
35	Major	10	3C	yes
36	Capt.	11	D	yes
37	Major	16	HS	no
38	Capt.	11	3C	no
39	Capt.	13	HS	yes
40	1st/Lt.	14	1C	no
41	Capt.	5	2C	yes
42	Lt/Col.	11	D	yes
43	Lt/Col.	25	HS	yes

SUMMARY OF BIOGRAPHICAL DATA USED IN ANALYSES

Subject Number	Rank	Years in Service	*Education	Instructor in Quartermaster School
44	1st/Lt.	4	D	yes
45	Capt.	10	2C	no
46	2nd/Lt.	1	D	no
47	Capt.	9	1C	yes
48	Capt.	12	HS	yes
49	Lt/Col.	19	3C	yes
50	2nd/Lt.	6	D	yes
51	Capt.	10	1C	yes
52	Capt.	12	2C	yes
53	1st/Lt.	11	HS	yes
54	Lt/Col.	15	1C	no
55	Lt/Col.	11	D	no

*Explanation of symbols used in education category.

1. High school education is abbreviated HS.
2. College degree granted is abbreviated D.
3. College attendance without earning a degree is abbreviated C. The number refers to the number of years completed.

PART V

Jastrow's Catalogue Used as Criterion
of Reporting Ability

JUDGES' RATINGS USED AS CRITERION OF REPORTING ABILITY

*Subject Number	Number Recognitions	Good Reporter	Poor Reporter	No Comment	**Coded Score
1	23	16	5	2	26
2	24	19	3	2	31
*3	4	2	2	0	15
*4	3	2	1	0	16
5	23	6	12	5	9
*6	2	2	0	0	17
7	17	11	3	3	23
8	37	29	5	3	39
9	21	8	8	5	15
10	35	27	3	5	39
11	23	8	13	2	10
12	10	4	4	2	15
13	23	10	8	5	17
14	27	14	5	8	24
15	24	15	4	5	26
*16	8	3	4	1	14
17	32	9	15	8	9
18	31	19	8	4	26
19	36	23	8	5	30
20	21	6	10	5	11
21	27	18	5	4	28
22	25	19	1	5	33
23	11	7	3	1	19
24	15	13	1	1	27
25	33	22	7	4	30
26	25	8	14	3	9
27	26	14	5	7	24
28	19	13	4	2	24
29	21	7	12	2	10
30	21	12	6	3	21
31	19	5	8	6	12
*32	8	4	2	2	17
33	10	3	5	2	13
34	23	13	6	4	22
35	26	23	1	2	37
36	33	20	10	3	25
37	25	11	7	7	19
*38	6	4	2	0	17
39	30	10	14	6	11
*40	8	4	2	2	17
41	17	5	9	3	11
42	35	30	3	2	42
43	32	22	6	4	31

JUDGES' RATINGS USED AS CRITERION OF REPORTING ABILITY

*Subject Number	Number Recognitions	Good Reporter	Poor Reporter	No Comment	**Coded Score
*44	7	3	4	0	14
45	12	4	4	4	15
*46	5	2	3	0	14
47	27	12	10	5	17
48	16	9	5	2	19
49	32	24	4	4	35
*50	9	6	1	2	20
51	20	15	3	2	27
52	11	4	7	0	12
53	30	14	13	3	16
54	20	6	10	4	11
55	24	19	1	4	33

*Those subjects with a star (*) before their number received fewer than ten recognitions and were excluded from the criterion group.

**The coded score, which represented each subject, consisted of the difference between the number of favorable and the number of unfavorable nominations, plus fifteen. This latter number assured no subject of a minus score.

APPENDIX F

Raw Data

Recall Scores by Indices of Reporting Ability for each Subject

Recall Scores by Picture for each Subject

Recognition Scores by Indices of Reporting Ability for each Subject

Recognition Scores by Picture for each Subject

RECALL SCORES BY INDICES OF REPORTING ABILITY
FOR EACH SUBJECT

Six Pictures Combined

Subject Number	*Length of Report	**Detail	Clothing Detail	Inference	Ambiguity
1	86	113	6	13	12
2	133	155	10	17	11
3	108	142	14	10	16
4	87	118	10	11	8
5	100	138	7	24	16
6	54	86	2	10	15
7	41	64	1	6	3
8	72	99	2	8	7
9	214	206	16	26	29
10	82	132	6	3	5
11	96	93	1	7	11
12	57	49	3	20	2
13	61	71	2	13	11
14	54	101	3	7	5
15	100	146	8	9	9
16	98	168	11	8	8
17	111	109	19	27	14
18	109	168	7	7	14
19	201	161	14	27	8
20	171	188	12	19	16
21	88	146	6	7	11
22	60	91	1	9	7
23	124	171	10	7	9
24	55	118	4	3	8
25	137	213	26	7	4
26	139	193	14	19	11
27	73	73	6	25	9
28	147	203	14	15	16
29	168	102	9	40	7
30	61	109	4	2	11
31	151	143	5	24	10
32	122	158	3	12	8
33	145	141	21	19	18
34	147	167	9	10	12
35	165	220	23	14	12
36	68	50	1	22	5
37	68	123	6	6	11
38	82	121	6	13	13
39	40	62	1	9	4
40	119	164	30	21	12
41	54	89	5	12	4

RECALL SCORES BY INDICES OF REPORTING ABILITY
FOR EACH SUBJECT

Subject Number	*Length of Report	**Detail	Clothing Detail	Inference	Ambiguity
42	70	111	2	8	6
43	98	136	9	11	6
44	99	180	4	10	9
45	61	70	3	11	1
46	102	142	13	19	8
47	108	141	11	3	13
48	50	89	1	7	6
49	122	156	11	4	11
50	43	93	6	4	7
51	173	211	20	18	16
52	123	152	7	10	6
53	83	137	6	3	9
54	98	148	13	8	17
55	52	81	5	5	12

*Length of report estimated by number of typewritten lines.
**Clothing detail is not counted in the detail score.

*RECALL SCORES BY PICTURE FOR EACH SUBJECT

Subject Number	Picture A	B	C	D	E	F	Total Score
1	25	23	24	27	21	24	144
2	27	26	35	28	42	35	193
3	33	26	30	32	26	35	182
4	15	27	35	31	17	22	147
5	28	37	32	32	25	31	185
6	23	19	18	17	18	18	113
7	14	16	7	12	9	16	74
8	10	20	21	26	23	16	116
9	41	44	35	59	53	45	277
10	19	30	25	24	27	21	146
11	13	24	19	18	18	20	112
12	17	15	7	13	11	11	74
13	8	22	18	17	20	12	97
14	13	21	23	28	16	15	116
15	21	36	31	35	26	23	172
16	24	40	30	38	36	27	195
17	28	36	17	33	20	35	169
18	29	41	31	35	35	25	196
19	25	52	35	43	30	25	210
20	39	42	45	40	26	43	235
21	22	26	33	31	33	25	170
22	13	17	24	19	13	22	108
23	26	35	38	34	34	30	197
24	22	25	24	21	17	24	133
25	38	50	39	41	34	48	250
26	46	51	48	27	26	39	237
27	16	15	14	25	22	21	113
28	41	48	47	36	36	40	248
29	23	27	18	36	24	30	158
30	11	29	26	22	22	16	126
31	30	31	27	30	31	33	182
32	32	32	30	39	25	25	183
33	41	37	35	35	21	30	199
34	18	35	45	35	26	39	198
35	65	45	24	39	49	47	269
36	14	17	13	16	4	14	78
37	12	19	34	23	29	26	143
38	27	21	36	25	19	25	153
39	14	7	12	19	7	17	76
40	39	35	32	40	36	45	227
41	13	15	20	26	17	19	110
42	18	23	16	33	18	19	127
43	21	31	26	30	30	24	162
44	26	35	42	28	34	38	203

*RECALL SCORES BY PICTURE FOR EACH SUBJECT

Subject Number	Picture A	B	C	D	E	F	Total Score
45	12	19	11	22	3	18	85
46	27	46	41	28	11	29	182
47	15	33	31	23	36	30	168
48	15	13	18	30	8	19	103
49	21	39	39	34	35	14	182
50	18	22	12	17	16	25	110
51	26	40	50	37	50	62	265
52	21	25	39	31	24	35	175
53	16	32	21	26	30	30	155
54	30	32	29	36	25	34	186
55	14	13	19	15	20	22	103

*The score for each picture consists of the sum of the scores for detail (excluding clothing), clothing detail, inference, and ambiguity.

RECOGNITION SCORES BY INDICES OF REPORTING ABILITY FOR EACH SUBJECT

Six Pictures Combined

Subject Number	*Don't Know	**Detail	Clothing Detail	Inference	Ambiguity
1	19	22	9	5	4
2	26	20	12	8	8
3	14	18	13	4	5
4	40	17	5	6	14
5	22	23	11	7	6
6	11	16	10	7	5
7	10	25	8	4	10
8	28	13	8	3	10
9	19	17	8	3	6
10	17	23	10	4	7
11	31	15	6	5	7
12	25	15	12	6	6
13	20	23	9	2	9
14	25	18	11	6	29
15	16	20	9	5	7
16	8	27	14	2	2
17	19	23	13	8	6
18	30	20	9	6	6
19	33	15	9	5	5
20	1	25	15	1	4
21	40	15	7	3	10
22	24	23	10	5	7
23	9	27	9	3	6
24	21	19	12	6	9
25	20	24	12	4	5
26	28	17	9	3	5
27	26	19	11	5	11
28	28	12	8	4	7
29	19	16	9	7	12
30	24	20	8	6	7
31	39	19	5	6	10
32	1	29	15	6	7
33	2	22	15	6	4
34	23	16	9	5	8
35	7	25	11	3	4
36	28	18	5	9	9
37	18	20	9	6	9
38	14	28	11	11	4
39	19	26	10	10	5
40	8	26	14	7	6
41	10	22	15	5	1

RECOGNITION SCORES BY INDICES
OF REPORTING ABILITY FOR EACH SUBJECT

Subject Number	*Don't Know	**Detail	Clothing Detail	Inference	Ambiguity
42	36	16	9	5	8
43	27	18	11	6	5
44	13	22	12	2	16
45	28	14	8	5	13
46	10	24	14	3	18
47	21	20	11	5	10
48	32	17	8	5	12
49	25	21	6	4	14
50	14	25	10	2	12
51	48	15	1	6	11
52	21	15	8	1	15
53	25	19	9	2	13
54	7	25	16	7	14
55	20	24	6	4	9

* Number of "don't know" responses was tabulated as a possible index.
** Excludes clothing detail.

*RECOGNITION SCORES BY PICTURE FOR EACH SUBJECT

Subject Number	Picture A	B	C	D	E	F	Total Score
1	6	8	8	11	11	11	55
2	3	8	7	9	10	12	49
3	8	8	9	10	10	11	56
4	5	7	8	5	8	2	35
5	8	8	8	7	12	11	54
6	8	6	6	8	9	10	47
7	4	7	7	13	11	10	52
8	6	4	10	6	9	6	41
9	9	6	8	6	10	10	49
10	7	8	7	10	12	11	55
11	8	8	5	8	8	5	42
12	8	7	7	7	9	10	48
13	7	8	8	7	13	11	54
14	6	8	6	8	9	10	47
15	7	7	9	9	8	10	50
16	13	9	7	14	14	13	70
17	8	7	6	13	7	14	55
18	7	8	6	11	11	7	50
19	7	8	7	8	10	6	46
20	12	8	10	13	13	12	68
21	8	6	7	7	11	3	42
22	6	6	6	10	12	14	54
23	10	7	6	13	13	11	60
24	8	5	8	8	11	9	49
25	7	8	7	14	10	14	60
26	12	9	8	8	9	5	51
27	5	8	9	10	10	5	47
28	9	5	7	4	8	9	42
29	7	6	6	5	8	7	39
30	6	6	7	12	10	7	48
31	7	5	6	6	9	8	41
32	5	10	9	14	11	14	63
33	13	9	7	13	10	8	60
34	7	5	4	8	11	10	45
35	10	9	8	13	10	12	62
36	5	3	6	12	7	5	38
37	3	9	7	8	11	9	47
38	10	7	5	14	9	11	56
39	6	8	7	13	10	10	54
40	8	6	7	12	12	15	60
41	11	10	9	11	11	12	64
42	8	8	7	4	9	9	45
43	9	9	8	7	10	8	41
44	14	8	9	9	13	7	60
45	5	5	7	8	10	7	42

*RECOGNITION SCORES BY PICTURE FOR EACH SUBJECT

Subject Number	Picture A	B	C	D	E	F	Total Score
46	10	10	10	12	11	12	65
47	8	6	7	9	12	6	48
48	6	6	7	6	9	10	44
49	7	8	8	10	12	4	49
50	10	7	8	8	13	11	57
51	2	6	7	5	10	3	33
52	11	4	7	6	10	10	48
53	5	7	8	12	11	9	52
54	13	6	4	11	11	15	60
55	7	4	9	6	12	9	47

*The score for each picture consists of the sum of detail (excluding clothing), clothing, correct inferences, and correct ambiguities.

VITA

Herbert Irving Abelson

1 Freeland Street, Worcester 3, Massachusetts

Degree to be conferred: Doctor of Philosophy, June 7, 1952

Date of birth: January 13, 1926

Place of birth: Worcester, Massachusetts

Secondary Education: South High School
Worcester, Massachusetts

Collegiate Institutions attended:

Clark University, 1943-1947, A.B., May 1947
University of Maryland, 1947-1949, M.A., June 1949

Publication:

An adaptation of the method of paired comparisons to the measurement of radio audience preferences. Paper read at meetings of Eastern Psychological Association, Springfield, Massachusetts, March, 1949.