

THIS REPORT HAS BEEN DELIMITED
AND CLEARED FOR PUBLIC RELEASE
UNDER DOD DIRECTIVE 5200.20 AND
NO RESTRICTIONS ARE IMPOSED UPON
ITS USE AND DISCLOSURE.

DISTRIBUTION STATEMENT A

APPROVED FOR PUBLIC RELEASE;
DISTRIBUTION UNLIMITED.

Armed Services Technical Information Agency

Because of our limited supply, you are requested to return this copy WHEN IT HAS SERVED YOUR PURPOSE so that it may be made available to other requesters. Your cooperation will be appreciated.

AD

37814

NOTICE: WHEN GOVERNMENT OR OTHER DRAWINGS, SPECIFICATIONS OR OTHER DATA ARE USED FOR ANY PURPOSE OTHER THAN IN CONNECTION WITH A DEFINITELY RELATED GOVERNMENT PROCUREMENT OPERATION, THE U. S. GOVERNMENT THEREBY INCURS NO RESPONSIBILITY, NOR ANY OBLIGATION WHATSOEVER; AND THE FACT THAT THE GOVERNMENT MAY HAVE FORMULATED, FURNISHED, OR IN ANY WAY SUPPLIED THE SAID DRAWINGS, SPECIFICATIONS, OR OTHER DATA IS NOT TO BE REGARDED BY IMPLICATION OR OTHERWISE AS IN ANY MANNER LICENSING THE HOLDER OR ANY OTHER PERSON OR CORPORATION, OR CONVEYING ANY RIGHTS OR PERMISSION TO MANUFACTURE, USE OR SELL ANY PATENTED INVENTION THAT MAY IN ANY WAY BE RELATED THERETO.

Reproduced by
DOCUMENT SERVICE CENTER
KNOTT BUILDING, DAYTON, 2, OHIO

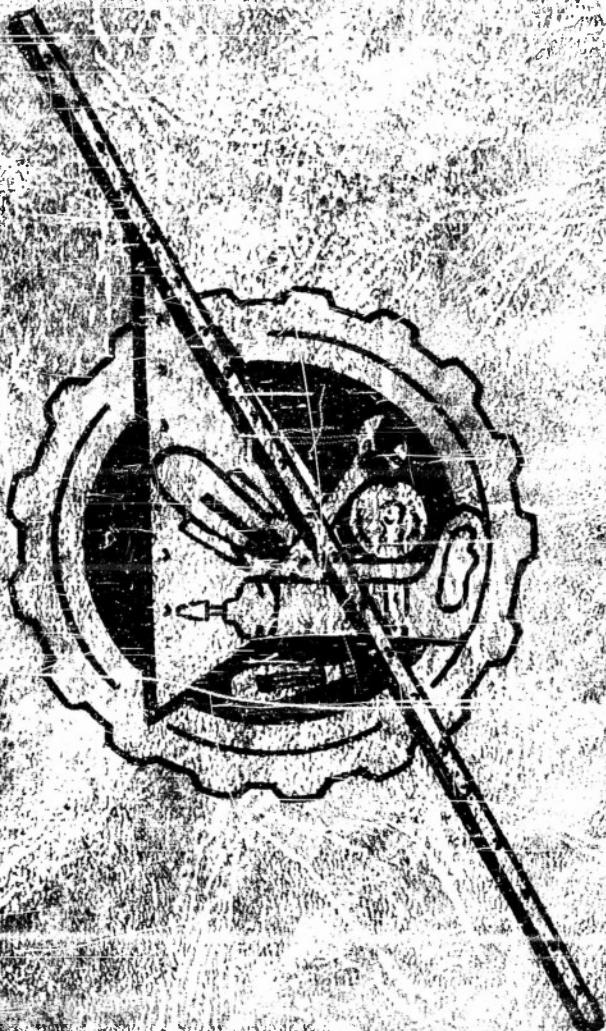
UNCLASSIFIED

AD NO. 37817

ASIA

ILLUSTRATION A-11

PSYCHOLOGICAL REPORT



NOAR Contract Research 616 (02)
Washington University
St. Louis, Mo

NAVAL AIR TECHNICAL TRAINING

ANNUAL REPORT

CONTRACT ONR 816 - 02

WILSE B. WEBB, PRINCIPLE INVESTIGATOR

WASHINGTON UNIVERSITY
SAINT LOUIS

DEPARTMENT OF PSYCHOLOGY

Dr. Denzel D. Smith
Head, Personnel and Training Branch
Office of Naval Research
Department of the Navy
Washington, D. C.

Dear Sir:

The attached three research notes are forwarded as a partial statement of our work in the Naval Air Technical Training research under NONR contract 816(02). These reports are exploratory studies conducted incidental to and concurrent with our major research effort, "A study of the activities of aviation machinist mates in fleet activities as a source for curriculum evaluation." This major report will be presented as a separate technical report.

In addition to the attached research notes and the technical report in preparation, we have been instrumental in the preparation of a classroom communicator and in a survey of research needs in Naval Air Technical Training. No formal report is made of these activities at the present time as they represent continuing efforts.

Finally, the understandings, the background knowledges, and the development of within service relationships cannot be expressed in an annual report but represent an important part of our activities during the year.

We have been well pleased with the cooperation extended by Naval Air Technical Training during the year. In particular, the aid furnished by Dr. G. D. Mayo has been invaluable. Your services in this contract are deeply appreciated.

Sincerely,

WILSE B. WEBB, Ph.D.
Head, Aviation Psychology
Laboratory

Research Note I

**Ratings of Instructional Proficiency: Student Ratings,
Self Ratings, and Supervisor Ratings**

Wilse B. Webb and Carson Y. Nolan

Office of Naval Research Contract NONR 816 (02)

RATINGS OF INSTRUCTIONAL PROFICIENCY: STUDENT RATINGS,
SELF RATINGS, AND SUPERVISOR RATINGS

INTRODUCTION

In the evaluation of teaching proficiency there are at least two classical methods of evaluation; student ratings and supervisor ratings. The present study is concerned with the relations between these evaluative procedures and a third evaluative source, the teacher's self evaluation of his own proficiency.

Many arguments have been put forth in favor of and against student ratings and supervisor (or peer) ratings: "the students are the consumer and hence must be the judge"; "the student is incapable of evaluating what he is learning at the time"; "the supervisor is the only one capable of evaluating, from a broad base of experience and understanding, what is to be taught"; "the supervisor does not have the opportunity to observe or his observations are distorted"; etc. These arguments have been too frequent to review here and too frequently merely argumentative. We would like to comment on the self rating as a source of evaluation because of its more infrequent use. Basically we feel that personal learning and improvement stems from an understanding of ones own adequacies and inadequacies. We feel that a self evaluation procedure serves to focus the individuals attention on his inadequacies and as such he will be motivated to attempt to correct them. Admissably this technique cannot serve as an administrative device since the man will distort these ratings for secondary purposes. However, in a non-threatening situation and in conjunction with other evaluative procedures, we feel that self ratings are a valuable adjunct to improving and evaluating teaching proficiency.

Procedure

A rating scale covering seven characteristics of importance to the teaching situation was administered to 51 instructors. The characteristics rated were interest of the instructor in his subject, his sympathetic attitude towards students, presentation of subject matter, sense of proportion and humor, self reliance and confidence, personal peculiarities, and personal appearance. Each of these characteristics was rated on a zero to ten point scale.*

The procedure for administering the ratings follows. The problem was explained at length to small groups of instructors. Particular emphasis was made on the point that the results of any of the ratings for any individual would not be disclosed. The instructors were then asked to rate themselves. After additional instruction about avoiding biasing the students, the proper

*This study was performed in the Naval Air Technical Training School at Jacksonville, Fla.

number of questionnaires were given the instructor. Since the students were not required to sign their names to their rating scale, the instructor was also given an envelope which was to be sealed after being filled with the completed student ratings. The instructor's name was then put on the envelope, and it was turned in. In all cases the students had at least eight hours of instruction before rating their teachers. The subject matter of the course was mathematics, physics, layout and hand tools.

A meeting of all the supervisors of the instructors used in the study was next called. The nature of the problem was again explained at great length and any questions answered. The supervisors were then asked to rate their instructors in the above form. In addition they were asked to rate the instructors on an Air Force forced choice rating form, the Instructors Description Form C which was developed by Highland & Bershire (1).

Biographical information about instructors was collected from various sources. These data included the GCT score of the instructors, the amount of formal education of the instructor, the number of years he had taught, and his enthusiasm for teaching as indicated by him on a 7 point scale. A number of product moment correlations were computed between these data and those derived from the two rating scales. These coefficients are listed in Table I.

TABLE I.

Coefficients of correlation between variable of study

Instructor self-rating and student rating	.62
Instructor self-rating and supervisor rating	.16
Student rating and supervisor rating.	.13
Instructor GCT score and instructor self-rating	.25
Instructor GCT score and student rating	.21
Instructor GCT score and supervisor rating.	.08
Instructor formal education and his own self-rating	-.23
Instructor formal education and student rating.	.18
Instructor formal education and supervisor rating	.08
Instructor's teaching experience and his own self-rating.	.05
Instructor's teaching experience and student rating	-.21
Instructor's teaching experience and supervisor rating.	-.13
Instructor's enthusiasm for teaching and his own self rating.	.25
Instructor's enthusiasm for teaching and student rating	.39
Instructor's enthusiasm for teaching and supervisor rating.	.00
Supervisor rating on Air Force form and on our form	.37
Supervisor rating on Air Force form and instructor self- rating	.20
Supervisor rating on Air Force form and student rating.	.09
Student ratings in stanine form and difference scores between self-ratings in stanine form and student ratings in stanine form	-.13

Note: With a sample population of 51 the five and one percent levels of significance are 0.27 and 0.35 respectively.

Discussion

Our main findings seem to be rather clear cut. There is a rather high relationship between the way the student views a teacher as a teacher and the way the teacher views himself. When the limits of reliability of group ratings and self ratings are taken into account this relationship is strikingly high (2). In other words, the teacher does have an idea of himself which is quite similar to the "consumers" idea when he is called upon to make such an evaluation. We feel that this insight can be used as a prime source of instructional improvement.

When we turn to supervisor ratings we find a different picture. There is little relationship between the students view of the teacher or the teacher's view of himself and the supervisors ratings. In fact it is difficult to tell what is the basis of the supervisor's rating since these ratings were not significantly correlated with the intelligence of the instructor (GCT), his experience in teaching, his level of schooling, or his enthusiasm for teaching. We can only conclude that the supervisor was rating on some factor or factors other than these which were valid estimates of the teaching ability of the individual or were random invalid intuitions.

Certain interesting points can be noted about the student ratings and the instructor ratings. Only one further correlation obtained was statistically significant. This was the positive correlation of .39 between the instructors enthusiasm for teaching and the students ratings of his teaching ability. This would clearly support the hoary but apparently sound generalization that one of the prime attributes of a good teacher is his desire to teach. We may further note that although not statistically significant, the next two highest correlations indicated that the more intelligent and the more educated instructors seem to be more self critical (correlations of -.25 and -.23 between the teachers self ratings and the GCT and level of schooling respectively).

In the particular situation it would seem possible to state that the GCT, the level of schooling or teaching experience (within the limits of the selected population) were not significant variables in the teaching situation.

The last correlation report in Table One was our greatest disappointment. We hypothesized that the greater the difference between the student rating and the self rating, the lower the student rating would be. This was based on the assumption that widely disparate (assigned or judged) roles between the student and the instructor would result in "psychological friction". As indicated by the -.13 correlation between these variables our hypothesis was not confirmed. Although the correlation is in the predicted direction it is far from statistically significant.

Summary and Recommendations

Fifty one instructors were rated by their students and by their supervisors on a "teaching proficiency" rating scale. In addition the instructors rated themselves on the same scale. It was found that the student ratings and the self ratings of the instructors were highly correlated. However, the supervisors ratings were uncorrelated with either of these ratings and, in fact, uncorrelated with any of the additional measures obtained (the instructors GCT, his level of schooling, his teaching experience, or his desire to teach). There was a tendency for the more intelligent instructors and those with more schooling to be more self critical. The instructors who expressed a greater desire to teach were rated as superior teachers by their students. The discrepancy between student ratings and the instructors ratings did not seem to be related to the judged proficiency of the teacher.

On the basis of these findings the following recommendations are made:

1. Systematic self ratings should be introduced as a potential source of self improvement in instructing. The Naval Air Technical Training supervisors of instructional training should develop or utilize available forms which include elements felt to play a role in instructional proficiency. These forms should be administered to the present instructional population to sensitize them to these factors judged to be critical in their role as instructors. It is felt that self improvement on these factors may result from such self evaluation.

2. Supervisors' ratings should not be used administratively and an immediate examination of the sources of supervisor ratings should be made.

3. The GCT, the level of schooling, and the instructional experience of instructors were found not to be related to instructional proficiency as judged by students or by the instructors themselves. Upon the basis of these findings, these factors should not be weighed heavily in the administration of the instructor program. This does not imply that they may not be used for selective purposes. Such an implication would be dependent upon further studies.

Bibliography

- (1) Highland, R. W. and Berkshire, J.R. A Methodological Study of forced-choice performance rating. Air Training Command, Human Resources Research Center, Bulletin 51-9.
- (2) Webb, W. B. and Stafford, Billy, Group Evaluations and self evaluations: The matter of reliability. Unpublished study available from senior author.

Research Note II

A Study of the Use of Open End Sentences
in the Naval Air Technical Training Program

Wilse B. Webb and Carson Y. Nolan

Office of Naval Research Contract NONR 816 (02)

A Study of the use of Open End Sentences
in the Naval Air Technical Training Program

Recently devised, the open ended sentence is one of the most flexible and potentially useful tools in the kit of the psychologist. The form of this tool and the assumptions underlying its use are gratifyingly simple. The subject is presented with a portion of a sentence (typically a noun and a verb), and he is asked to complete the sentence. The experimenter may leave the sentence or the situation quite unstructured by presenting such incomplete sentences as "Most people" or "I wish", etc. The experimenter may increase the structuring of the problem by instructions to the subjects or by making the sentences more directly pertinent to his needs. For example, with the highly unstructured sentences mentioned above, the experimenter may say, "Fill these sentences out in relation to this class", or "Fill these sentences out in relation to your childhood". He may use such sentences as "The instructor" or "The tests in this course" or "Your mother" depending on his particular area of interest. This possibility of directional structuring of the sentences permit the use of this technique in an unlimited number of specific problem areas.

The assumptions underlying the use of such sentences are common to all of the "projective" techniques being widely used in psychology today. Since the structure of the response is not inherent in the question submitted to the subject, the structure given in response to the sentence must necessarily reflect the subject rather than the experimenter. That is, the completion of the sentences by the subject must necessarily tell something about the subject since he alone is the source of the response. The advantages of this position are numerous and have been thoroughly reviewed in various publications. The simplest of these advantages for our problem is that qualitative responses can be obtained which are unthought of or unknown to the experimenter.

The major difficulties in the use of the projective approach typically lie in the inability to score or classify the responses which are obtained. Further, there is always the question as to whether these responses reflect anything consistent or meaningful about the subject, or whether they are just random thoughts that happen to pop into the subject's mind at the time.

The present project was initiated to answer several questions:

- 1) Can the open end sentence be used effectively in the Naval Technical Training situation?
- 2) Can the data be scored?
- 3) Can useful qualitative information be obtained in this manner?
- 4) Are there any consistencies in the responses of individuals elicited by this method which may indicate other than transitory opinions?

If the last three questions could be answered affirmatively, it was believed that the open ended question would be a worthwhile device to use in the Technical Training program and would yield data which would be superior to other more direct and less systematic methods of evaluation. This would be particularly true in the area of class room training toward which the questions were particularly structured in this study.

Experimental Procedure for Open End Study

The questionnaire used in the study consisted of twenty-one open ended sentences which are listed below.

1. The instructors
2. This school
3. I like
4. I don't like
5. I feel that I learned
6. What bores me
7. The tests
8. The rate at which the material is presented
9. I feel that questions
10. The set-back system
11. Other people in my class
12. What worries me
13. The type students
14. I hate
15. I wish
16. Most people
17. The biggest trouble
18. The Navy
19. The best
20. Very few
21. Nearly all

Instructions for the questionnaire were as follows: "Using the subjects below complete the sentences. Express your real feeling. Work rapidly. Complete every one. Be sure and make a complete sentence."

The questionnaires were administered to thirty students in the Aviation Machinist's Mates Class "A" School. The student was not required to sign the questionnaire. Instead he was asked to place some "alias" at the head of the sheet and to remember this "alias". Two weeks later the questionnaire was re-administered to the same group. They were asked to identify their papers with the "alias" they had used previously.

The collected data were then rated in two ways. Those open-end questions the stems of which contained a definite subject (1,2,5,7,8,9,10,11,13,16,18) were rated by two raters as expressing a positive attitude towards the subject of the sentence, a neutral attitude, or a negative attitude. Those open-end questions not having a definite subject in the stem (3,4,6,12,14,15,17,19,20,21) were placed in one of the following three categories. Category I

included answers relating feeling directly associated with technical training; category II included answers relating feelings associated with life generally; and category III was reserved for evasive answers.

Following independent ratings, by two raters, the rater and test-retest reliabilities were computed. A qualitative analysis of the results were also made.

Results

When computed using the product-moment formula the reliability between raters was .89. The test-retest reliabilities for individuals was .69. Algebraic summations of the +, 0, and - scores on the subject matter stems were used in this computation.

Below are listed each of the subjects for which attitudes were rated as positive, neutral, or negative and the percent of the sample rated under each attitude. Because of the small N (30), further qualitative analysis was not made on this data.

Subject	Positive Attitude	Neutral Attitude	Negative Attitude
1. instructors	80%	3%	17%
2. this school	63	7	30
5. I feel that I learned	83	4	13
7. the tests	45	13	42
8. rate of presentation	40	13	47
9. questions	62	4	34
10. set-backs	61	10	29
11. other people	64	25	11
13. type students	36	43	21
16. most people	69	7	24
18. the Navy	37	13	50

When the unstructured open-end sentences (3,4,6,12,14,15,17,19,20,21) were classified as indicating position (3,19), neutral (15,20,21), and negative (4,6,12,14,17) attitudes the following subjects were mentioned by more than five people.*

A. Positive attitudes were expressed towards:

1. having more work on the line (16 people)
2. having more work in trouble shooting (6)
3. the instructors (8)

*It is to be expected in this instance that a considerable proportion of the "unstructured" responses would be "structured" toward the instructional program because of the "structure" of the other questions and the circumstances of administration (the class room).

B. Neutral attitudes were expressed towards:

1. food in the mess halls (5)
2. instructors (7)

C. Negative attitudes were expressed towards:

1. night school (8)
2. the kind of tests used (15)
3. the guards at the gates (5)
4. attending classes (22)
5. the food in the mess halls (21)
6. petty regulations (9)
7. their classmates (5)
8. getting the desired next assignment (7)

Summary

In summary, it was found that 21 opened ended sentences (eleven of which were pointed at a definite subject matter and ten were completely unstructured as to subject matter) could be scored reliably in independent judges as to whether the statement given was an expression of a "positive" (or favorable) attitude, a "neutral" attitude, or a "negative" (or unfavorable) attitude. We further found that these attitudes were consistent from individual to individual in a retesting situation. Finally, we found the specific answers given to the completely unstructured items yield fruitful information about the general training program itself.

Recommendations

1. The open ended sentence is a flexible and simply constructed devise. Since this technique can be reliably quantified it is recommended that it be more widely used as an evaluation procedure for instructors, programs, or more generalized morale questions where positive, or negative attitudes of the subjects are considered critical.

2. Further, where it is desirable to explore the qualitative "positive" or "negative" factors with individual instructors or with programs the open-ended sentence approach is recommended.

Research Note III

The Effect of Accumulative Retesting on
Total Retention of Course Material

by

Wilse B. Webb

Project Director

C. Y. Nolan

Research Psychologist

ONR Contract 816 (02)

The effect of Accumulative Retesting on Total Retention of Course Material

Introduction

If course content is broken into several relatively discrete units which extend over a number of weeks, and if recall is not elicited at later periods for material appearing earlier in the course, a substantial amount of forgetting of the earlier material may occur before completion of the course by the students. This highly probable state of affairs is quite unfortunate for a training program. Regardless of the amount of learning that is originally developed unless the process of forgetting (or loss of this learning) is attended to much of our efforts are wasted. This project represents an attempt to introduce and evaluate a classical technique of reducing forgetting in the school learning situation . . . the method of successive retesting.

Successive recalls are the analogue of reviews, informal study and teaching, and for pedagogical purposes may take the form of a test. Thus, tests represent, in one sense, a relearning period. Spitzer (4) has demonstrated that a multiple choice test is an effective form of review for sixth grade pupils. Pupils who were given frequent retests over a period of 63 days, made significantly higher final scores than did those who had been given no retests. These results are supported by those of Spencer (3) and a number of other studies.

The Experimental Design

It was decided to test the hypothesis that progressive retesting would increase students' final comprehensive scores in the Aviation Machinists Mates Class "A" School at the Naval Air Technical Training Center, Memphis, Tennessee. The curriculum of this school is divided into nine areas over which the student is tested as the areas are completed. He is not retested over this material until his final comprehensive examination. At the end of the course, the student is given a comprehensive test over all nine areas.

Three experimental groups were set up to test the hypothesis and were identified by the letters, A, B, and C. Group C was the control group and, as such, was allowed to go through the course in the normal fashion. Group A was told that it would be progressively tested over all previous material. These students would receive the normal area examination and in addition they would receive additional questions over all the previous material which were cumulatively added to the retest items. Group B was told that it would also be responsible for all previous material. This Group was retested, however, on only the material in Phase I and Phase II of the course at the time of each subsequent area test. Groups A and B represented an experimental variation to determine whether any obtained facilitation on the final examination could be attributed to the retesting procedure per se (in the form of a motivational device) or could be attributed to this motivational factor plus

the actual retesting (review) of the specific material of each phase.

In detail, Group A had approximately 50 questions added to their Phase III examination on the material of Phases I and II; at the end of Phase IV approximately 50 questions were added to their phase examination, one-half of which were concerned with Phases I and II and one-half concerned with Phase III. In other words the added retesting questions of any one phase included all previous phases in a proportional amount. Group B had the same number of questions added to the Phase examinations but the retesting material included material only concerned with Phases I and II. In each of the retestings the questions were different from the previously used retest questions or the questions on the final comprehensive examination.

Both Groups A and B were told that the additional questions would be counted as part of their total test scores. Data on 134 individuals were accumulated in Group A, 129 in Group B, and 127 in Group C.

Results of the Study

The mean scores for Groups A, B, and C on the final comprehensive examination together with their standard deviations are listed in Table 1 below. It is apparent that no significant differences exist between these means.

Table 1

Means and Sigmas for Final Comprehensive Scores

	<u>Means</u> Standard Score	Raw score*	<u>S.D.</u> Standard Score
Group A	74.77	103.2	11.54
Group B	74.56	103.5	11.42
Group C	74.14	102.9	11.18

*With a total of 150 possible

An analysis of the percentage of correct answers to test items covering Phase I and Phase II in all retests for Groups A and B was made and is presented below. No significant differences exist between the two groups. In addition, are presented chi squares between these frequencies and those that would normally occur by chance with the use of four multiple choice items.

Table 2

Percentage of Correct Items Covering Phases
I and II in the Retests for Groups A and B
and the Chi Squares of the Frequencies when
Compared with Chance Expectancies

Retest	Group A	Chi Square	Group B	Chi Square
1	22.3	.52	21.6	.83
2	25.7	.17	26.1	.09
3	33.3	5.06	21.9	.68
4	15.8	5.38	25.1	.09
5	24.0	.07	20.1	1.57
6	27.5	.45	29.1	1.07
7	23.4	.18	20.1	1.57
8	26.4	.15	26.5	.14
Mean:	24.9		23.9	

Note: With one degree of freedom, a Chi Square must be greater than 5.412 to be significant at the 2 per cent level.

Discussion of Results

The results of this study appear to indicate that progressive retesting over previously presented material during the sequence of a course of study does not facilitate the total amount of material retained. These results are blantly contradictory of theory, previous experimental findings, and common sense. In such a case it would seem wiser to either not report the results, to question the design and the experimental control of the experiment, or to attempt to learn something from the data on hand. Any of these courses are unpleasant for the experimenter but we have chosen the latter two alternatives.

The question of the experimental design. In retrospect, we are still satisfied with the basic design of the experiment. We would change only one feature. We would have the retest reviewed after each administration in order that each testing would actually constitute a review rather than assuming that each test item "forced" an implicit review on the part of the student. It is quite conceivable that "knowledge of results," which has been shown to be requisite for learning, is equally necessary for relearning.

The question of the situation. Certain questions may be raised about the situation which if answered affirmatively would make for circumstances in which the hypotheses of facilitation could not be expected to be true.

(1) Was the original learning at such a level or the rate of forgetting such that the motivational factors of retesting resulted in no recall? If no recall occurred or retention was at such a low level that a question could not be judged as true or false on a retest and in addition the students were not told which answers were correct, obviously no relearning could occur. There is evidence to support a hypothesis of a very low level of learning or a hypothesis of extremely rapid forgetting when we note that the percentage of correct answers on Phase I and II material on the first retesting did not differ significantly from answers which could have been obtained from chance guessing (Table 2). This is true for the first retesting. The limitations of obtaining increases in retention would be true also for subsequent retestings (also indicated by the chance results on these retests).

(2) Was sufficient retention present but the retesting situation failed to serve as a motivating situation? If this was true the retesting would serve neither as a source of stimulation for implicit rehearsal or for the elicitation of correct responses which in themselves would be a review or learning trial. Again, looking at the actual percentage of correct responses on the retesting situations, we could support such a hypothesis. If this hypothesis were true, the same effects as indicated by our first question would be true and we could not expect an increase in retention.

According to the educational advisor of the Aviation Machinist Mates' School, this possibility of low motivation to respond should be weighed heavily. Although no direct evidence can be presented, the educational advisor has indicated that several incidental factors indicate that the trainees obtained information that these test results were merely experimental and were not part of their records. As such, the trainees made little or no attempt to give highly motivated responses to the testing situation.

(3) Did the retests fail to measure anything relevant to the original learning and its retention? It is quite obvious that if questions were asked which were quite irrelevant to the original learning and its retention, they would in no way serve as a review of this original learning. Similarly, if the final exam was unrelated to what was actually known by the student, it could not possibly be facilitated by reviews of any type.

From the data available it is impossible to select the most admissible of the hypotheses given above. All three of these would result in the chance figures obtained during retesting. It is possible, however, to view any of the hypotheses with alarm from a training point of view.

The Meaning of the Results. It seems possible to reject these results as an adequate test of the facilitatory effect of retesting, but it is clearly not possible to reject these results as presenting a considerable problem in regard to retention of material learned in the Naval Air Technical Training program. Our results indicate that retesting on material some two weeks after

learning yields chance results or results which would be obtained by individuals who had never had such training. Similarly, the results of the final comprehensive exam gave raw scores of approximately 69 per cent correct information over their program of training. If we take into account that a part of this raw score would be achieved on a mere guessing basis this means that only about 58% of what the men were taught is revealed on final examination over this material.* Certainly, these are figures to cause concern. It seems necessary that we attempt to answer whether this is a problem of low original learning, of rapid forgetting, of low motivation to produce what is learned, or an inadequacy of the testing itself. It is proposed that a study which will attempt to answer these questions should be immediately initiated.

Summary and Conclusions

A study of the effect of retesting on retention of material was performed. Three groups were used, involving about 130 cases each. One group received no retesting up to a final comprehensive exam. Another group received retesting only on approximately the first eighth of the course in eight sessions. A third group received accumulative retesting on all of the previous materials in eight sessions. The results indicated no significant effect between these treatments on the overall retention of the course material.

An examination of the level of retention on the retests and the level of the retention on the final comprehensive exam leads us to question the possibilities of obtaining such an effect rather than rejecting the possibility that retesting over material is not effective in increasing retention.

Most critically, the levels of retention indicated by our testings seems to demand that a more general evaluation of the learning, the retention, the motivation, and the testing in this area be performed.

*Like the retention scores on the interim test, the 58% retention on the final exam presents difficulties in interpretation. The score may represent low motivation, low retention or, in this case, may represent the fact that the items were selected for discrimination purposes and, hence, reflect item selection rather than an over-all evaluation of the percent of material retained.

Armed Services Technical Information Agency

Because of our limited supply, you are requested to return this copy WHEN IT HAS SERVED YOUR PURPOSE so that it may be made available to other requesters. Your cooperation will be appreciated.

AD 37814

NOTICE: WHEN GOVERNMENT OR OTHER DRAWINGS, SPECIFICATIONS OR OTHER DATA ARE USED FOR ANY PURPOSE OTHER THAN IN CONNECTION WITH A DEFINITELY RELATED GOVERNMENT PROCUREMENT OPERATION, THE U. S. GOVERNMENT THEREBY INCURS NO RESPONSIBILITY, NOR ANY OBLIGATION WHATSOEVER; AND THE FACT THAT THE GOVERNMENT MAY HAVE FORMULATED, FURNISHED, OR IN ANY WAY SUPPLIED THE SAID DRAWINGS, SPECIFICATIONS, OR OTHER DATA IS NOT TO BE REGARDED BY IMPLICATION OR OTHERWISE AS IN ANY MANNER LICENSING THE HOLDER OR ANY OTHER PERSON OR CORPORATION, OR CONVEYING ANY RIGHTS OR PERMISSION TO MANUFACTURE, USE OR SELL ANY PATENTED INVENTION THAT MAY IN ANY WAY BE RELATED THERETO.

**Reproduced by
DOCUMENT SERVICE CENTER
KNOTT BUILDING, DAYTON, 2, OHIO**

UNCLASSIFIED