Research Report 1341

Use of Self Assessments in Estimating Levels of Skill Retention

Joel D. Schendel, John C. Morey,
M. Janell Granier, and Sid Hall

ARI Field Unit at Fort Benning, Georgia
Training Research Laboratory

U. S. Army
Research Institute for the Behavioral and Social Sciences

September 1983

U. S. ARMY RESEARCH INSTITUTE

FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Technical Director

L. NEALE COSBY
Colonel, IN
Commander

Technical review by:

Billy L. Burnside
Stanley F. Bolin
Joseph D. Hagman

NOTICES

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: Research Report 1341	2. GOVT ACCESSION NO.: A141 Ø42
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): USE OF SELF ASSESSMENTS IN ESTIMATING LEVELS OF SKILL RETENTION	5. TYPE OF REPORT & PERIOD COVERED: Research Report
	6. PERFORMING ORG. REPORT NUMBER: --
7. AUTHOR(s): Joel D. Schendel, John C. Morey, M. Janell Granier, and Sid Hall	8. CONTRACT OR GRANT NUMBER(s): --
9. PERFORMING ORGANIZATION NAME AND ADDRESS: U.S. Army Research Institute for the Behavioral and Social Sciences, 5001 Eisenhower Avenue, Alexandria, VA 22333	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: Work Unit ID 5710 2Q263743A794
11. CONTROLLING OFFICE NAME AND ADDRESS: U.S. Army Research Institute for the Behavioral and Social Sciences, 5001 Eisenhower Avenue, Alexandria, VA 22333	12. REPORT DATE: September 1983
	13. NUMBER OF PAGES: 25
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): --	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE: --

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

--

18. SUPPLEMENTARY NOTES

--

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Retention	Marksmanship
Refresher training	M16A1 rifle training
Self assessment	Questionnaire design

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The focus of this research was on one approach toward predicting task retention and refresher training requirements--self assessment. Prior to zeroing their weapons for annual M16A1 rifle qualification (record fire), 153 permanent party soldiers completed a questionnaire designed to collect information on their previous marksmanship experiences as well as on their beliefs about how they were going to shoot at record fire. No special instructions were provided about the meanings of questions or possible responses. Questionnaire items then were correlated with rifle qualification scores. (Continued)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

 SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Item 20 (Continued)

based on 40 target exposures. Predicted scores accounted for about 10% of the variance associated with record fire scores. However, for those soldiers whose confidence in the accuracy of their predictions exceeded 90%, predicted scores accounted for 25% of the record fire score variance. The best predictors of record fire performance were remembered most recent record fire performance and predicted performance. Subjects, generally, overestimated their actual performance and were biased heavily toward predicting success. The usefulness of the self-assessment approach to skill retention estimation is discussed together with practical suggestions for refining the precision of questionnaire techniques.

Research Report 1341

Use of Self Assessments in Estimating Levels of Skill Retention

Joel D. Schendel, John C. Morey,
M. Janell Granier, and Sid Hall

Submitted by
Seward Smith, Chief
ARI Field Unit at Fort Benning, Georgia

Approved as technically adequate
and submitted for publication by
Harold F. O'Neil, Jr., Director
Training Research Laboratory

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES
5001 Eisenhower Avenue, Alexandria, Virginia 22333

Office, Deputy Chief of Staff for Personnel
Department of the Army

September 1983

Army Project Number
2Q263743A794

Education and Training

ARI Research Reports and Technical Reports are intended for sponsors of R&D tasks and for other research and military agencies. Any findings ready for implementation at the time of publication are presented in the last part of the Brief. Upon completion of a major phase of the task, formal recommendations for official action normally are conveyed to appropriate military agencies by briefing or Disposition Form.

FOREWORD

The Fort Benning Field Unit of the Army Research Institute for the Behavioral and Social Sciences (ARI) has conducted numerous experiments on specific topics related to skill acquisition, retention, and transfer. The research reported here was performed in response to the question, "How often does refresher training need to be provided to sustain skills continuously at high levels?" It represents one attempt to provide the Army training community an answer to this question and was intended to complement larger, related efforts by ARI's Training Technical Area. Ms. Granier and Mr. Hall were serving with ARI, Fort Benning Field Unit, under the auspices of Auburn University's cooperative education program when this research was conducted.

EDGAR M. JOHNSON
Technical Director

ACKNOWLEDGEMENTS

The support of soldiers assigned to the 1st and 2d Infantry Training Brigades and Infantry Training Group at Fort Benning, Ga is gratefully acknowledged. The helpful comments of Dr. Billy L. Burnside, Dr. Stanley F. Bolin, and Dr. Joseph D. Hagman also warrant special recognition.

USE OF SELF ASSESSMENTS IN ESTIMATING LEVELS OF SKILL RETENTION

EXECUTIVE SUMMARY

Requirement:

This research was conducted in response to the question, "How often does refresher training need to be provided to sustain skills continuously at high levels?" The focus was on one approach toward predicting task retention and refresher training requirements--self assessment. The emphasis was on identifying questionnaire techniques for improving the accuracy of these assessments.

Procedure:

Prior to zeroing their weapons for annual M16A1 rifle qualification (record fire), 153 permanent party soldiers volunteered to complete a questionnaire. This questionnaire was designed to collect information on subjects' previous marksmanship training experiences as well as to assess their beliefs about how they were going to shoot at record fire. Subjects completed the questionnaire in groups of 15 to 60 without special instructions about meanings of questions or possible responses. Record fire involved firing a total of 40 shots at silhouette targets seen at 50, 100, 150, 200, 250, and 300 meters. The first half of record fire was shot from the foxhole supported postion; the second half from the prone unsupported position. Targets fell when hit. Scoring was performed by independent support personnel.

Findings:

1. Subjects' self assessments accounted for about 10% of the variance associated with record fire scores.

2. Reported experience outside the Army firing a rifle or a shotgun failed to correlate with performance at record fire. This also was true for reported experience firing a rifle (other than the M16A1) or a shotgun since the last record fire and for reported training on the M16A1 since the last record fire. Nevertheless, those reporting such experiences or training thought they would shoot better at record fire than subjects who did not.

3. One of the best predictors of record fire performance was his (her) remembered shooting classification from his (her) most recent record fire. Remembered shooting classifications and scores also correlated highly with predicted shooting classifications and scores.

4. Subjects who estimated their chances of accurate prediction at 90% or 100%, generally, were most accurate in their predictions. These subjects' predicted scores accounted for 25% of the variance associated with their record fire scores.

5. Subjects who reported firing record fire in the last 12 months appeared more accurate in their self assessments than subjects who reported not firing record fire as recently.

6. Seventy-five percent of the subjects predicted they would hit more targets than they actually hit.

7. Subjects were biased heavily toward predicting success. Nevertheless, those who predicted failure were quite accurate.

Utilization of Findings:

The self-assessment methodology may serve to strengthen and refine data derived from other, more generalized approaches toward predicting task retention and refresher training requirements. In addition, findings reported here suggest means for improving the accuracy of self-assessment questionnaires.

USE OF SELF ASSESSMENTS IN ESTIMATING LEVELS OF SKILL RETENTION

CONTENTS

	Page
INTRODUCTION	1
METHOD	3
Subjects	3
Materials and Procedure	3
RESULTS	4
Overall	4
Questionnaire Data Corrected For Internal Consistency	5
DISCUSSION	8
REFERENCES	11
APPENDIX A. SELF-ASSESSMENT QUESTIONNAIRE	A-1

USE OF SELF ASSESSMENTS IN ESTIMATING LEVELS OF SKILL RETENTION

INTRODUCTION

Most Army jobs require that soldiers achieve and maintain proficiency on scores of tasks. Because of resource constraints it is impossible to train every soldier on every task to the degree necessary even to minimize forgetting. As a result, some amount of periodic refresher training must be provided. This much is clear. What is not clear is how frequently individuals require refresher training to sustain proficiency on particular tasks. As Schendel and Hagman (1982) indicate, if time intervals between successive sessions are too long, then performance may fall below acceptable levels and entail considerable risk. Emergencies may arise requiring corrective action before an individual has had a chance to retrain. If time intervals between sessions are too short, then administrative costs necessarily are inflated.

What can be done to remedy this problem? Rose and Ford (1982) have focused their efforts on the development of a task classification system designed to aid in the prediction of refresher training requirements. Their idea is to group tasks into a few distinct categories. This grouping would be based on the presence of task characteristics or features known to influence retention (e.g., number of steps comprising a task) or refresher training requirements (e.g., task criticality). Efforts also are underway to develop rules to account for the effects of initial ability and training, both of which may influence performance (e.g., Schendel, Shields, & Katz, 1978), although neither is task specific.

One drawback with this approach is the time and expense involved in developing a manageable categorization system. Not only must category variables be defined and tested, but rules must be developed which are simple enough to enable users to categorize tasks quickly and accurately. A second drawback is that the results of most categorization systems tend to be overly general. Assumptions must be made when categorizing tasks which make it virtually impossible to deal with particulars—a particular individual (or group of individuals) trained in a particular way on a particular task. As a result, Rose and Ford's (1982) classification system may not be much help when it comes to trying to identify soldiers who require training on particular tasks.

A different approach is being taken by Rigg and Gray (1981). This approach involves the use of a stochastic model to predict training and refresher training requirements. This approach appears effective in predicting soldiers' refresher training requirements, at least on cognitively oriented and procedural-motor tasks (e.g., M16A1 maintenance), but it has at least two potential drawbacks in its application: (a) Trial 1 error data are required for all tasks for which training/refresher training predictions are to be generated. Costs involved in collecting and analyzing these data must be considered when computing the cost of this model to the Army. (Note, too, that Trial 1 error data may have to be collected more than once per task, assuming Trial 1 errors would change as conditions for training change (e.g., cutbacks in training time, reductions in training ammunition, introduction of a new training device). (b) Special attempts must be made during training and refresher training to ensure that everyone is trained to the same criterion. Predictions are inaccurate if criterion performance is not achieved or is not achievable.

Self assessment represents another possible approach toward managing refresher training. Under this approach, priorities may be assigned to tasks to be trained (managerial decision). Unit commanders then would determine where they are weakest, next weakest, etc., by sampling soldiers' estimated skill levels on these tasks. Priorities would be weighed against weaknesses as well as resources (time, travel requirements, equipment constraints, etc.) and unit refresher training schedules developed accordingly. Of course, self assessments need not be used alone. They may be used to identify soldiers requiring training on particular tasks and, consequently, may be used to strengthen and refine predictions derived from other, more generalized approaches (e.g., Rose & Ford, 1982; Rigg & Gray, 1981).

The advantages of the self-assessment approach are that it is simple and economical--self assessments can be collected using only a questionnaire. Self assessments also can be very accurate. For example, Levine, Flory, and Ash (1977) found significant positive correlations between subjects' ratings of their abilities and written test scores in areas such as spelling, grammar, reading, and arithmetic. They also found that self assessments of typing speed correlated .60 and above with results of a standardized typing test. The main disadvantage of self assessments is that they are subject to distortion, both intentional and unintentional. For example, the stereotype of a "good soldier" may influence self assessments. As a result, soldiers may be reluctant to admit needing training on tasks which they should know. Alternatively, soldiers may unintentionally overestimate their abilities to perform a task as a result of their overconfidence. Several methods for handling these problems are discussed later in this report. Many of these methods also have been covered by Burnside (1982) in his review of the literature on self assessment.

This research focuses on the use of self assessments as a means for estimating soldiers' retention levels. Permanent party soldiers first completed a questionnaire designed, in part, to assess their beliefs about how they were going to perform during their annual M16A1 rifle qualification (hereafter referred to as "record fire"). These soldiers then fired record fire, after which their predicted and actual record fire scores were correlated. Discussion centers around considerations for the design of questionnaires that may serve to enhance the accuracy of self assessments.

METHOD

Subjects

The subjects were 153 male ($n = 147$) and female ($n = 6$) soldiers assigned to the 1st and 2d Infantry Training Brigades and Infantry Training Group at Fort Benning, GA, who fired record fire between 26 and 27 March and 4 and 5 June 1982 and who volunteered to participate in this research. Subjects completed the questionnaire in groups ranging in size from approximately 15 to 60. Reported times since last record fire ranged from 1 to 60 months ($n = 147$, $M = 12.80$, $SD = 10.93$).

Materials and Procedure

Questionnaire data. Prior to firing record fire, soldiers reported to a 25-meter range to zero their weapons. Zeroing involves firing series of three-round shot groups at a specially designed 25-meter target and adjusting the rifle's sights until point of bullet impact coincides with point of aim. Prior to zeroing, soldiers assembled in bleachers to receive a safety briefing. After soldiers had assembled in the bleachers but before they had received this briefing, they were informed about the purpose of this research. Those who volunteered to participate completed the questionnaire found in Appendix A. Subjects were told merely to answer the questions as accurately as possible. No special instructions were provided about the meanings of questions nor were possible responses discussed.

Performance data. After zeroing, all soldiers were transported to an M16A1 rifle qualification range. On arriving, each subject was assigned randomly to one of eight firing lanes. All scoring was done by independent support personnel. These personnel were fully informed about scoring procedures and the purpose of this research prior to the onset of testing. In addition, an experimenter and numerous range personnel were available to assist in scorekeeping and to answer any questions arising during the course of testing.

The course of fire consisted of 14 "F-type" silhouettes and 26 "E-type" silhouettes. F-type silhouettes are designed to appear like the head and shoulders of a man and were seen at 50 ($n = 5$) and 100 ($n = 9$) meters. E-type silhouettes, designed to appear like the head and torso of a man, appeared at 150 ($n = 10$), 200 ($n = 8$), 250 ($n = 5$), and 300 ($n = 3$) meters. The first half of this course was shot from the foxhole supported position; the second half, from the prone unsupported position. Targets fell when hit. Record fire scores associated with particular shooting classifications are as follows: 0 to 22--Unqualified; 23 to 29--Marksman; 30 to 35--Sharpshooter; 36 to 40--Expert.

RESULTS

Overall

Pearson correlation coefficients (rs) were computed on all subjects' questionnaire and record fire data. Many more rs were computed than are reported here. The reasons for presenting the ones shown should become apparent in the discussion section.

Subjects' predicted and actual record fire scores correlated positively, $r = .26$ ($n = 150$, $p < .01$). A similar result was obtained when computing the r between subjects' predicted shooting classifications--expert, sharpshooter, marksman, or unqualified--and their actual classifications, $r = .25$ ($n = 152$, $p < .01$). While these rs appear low, they were substantially stronger than others that we expected might be quite strong. In particular, reported experience outside the Army firing a rifle or a shotgun failed to correlate with performance at record fire, $r = -.10$ ($n = 152$, $p > .05$). This also was true for reported experience firing a rifle (other than the M16A1) or a shotgun since the last record fire, $r = -.05$ ($n = 153$, $p > .05$). Even reported training (versus no training) on the M16A1 since the last record fire did not correlate with record fire performance, $r = -.08$, ($n = 151$, $p > .05$). This latter result is understandable. Only 39 (25.5%) soldiers reported receiving any form of training on the M16A1 since last firing record fire. Only 20 (13.0%) of these reported firing any live rounds, and only 6 (4%) reported spending any time dry firing or firing with blanks. In passing, it should be noted that while reported experience outside the Army firing a rifle or shotgun did not correlate with actual record fire performance, it did correlate with predicted record fire performance. In general, those reporting outside experience thought they would hit more targets, $r = .31$ ($n = 149$, $p < .01$), and receive a higher shooting classification, $r = .30$ ($n = 151$, $p < .01$), than those who did not. These results also held true for those who reported firing a rifle (other than the M16A1) or shotgun since the last record fire $r = .26$ ($n = 150$, $p < .01$); $r = .32$ ($n = 152$, $p < .01$), and, to a more limited degree, for those reporting some form of M16A1 training since the last record fire, $r = .19$ ($n = 148$, $p < .05$); $r = .19$ ($n = 150$, $p < .05$).

One of the best predictors of how well a soldier was going to shoot at record fire was his (her) remembered shooting classification from his (her) most recent record fire, $r = .38$ ($n = 133$), $p < .01$). Remembered scores from the most recent record fire also correlated with record fire performance, $r = .29$ ($n = 121$, $p < .01$). Also, remembered classifications and scores related highly to predicted classifications and scores. Remembered classifications correlated $r = .65$ ($n = 132$, $p < .01$) with predicted classifications and $r = .53$ ($n = 130$, $p < .01$) with predicted scores; remembered scores correlated $r = .67$ ($n = 120$, $p < .01$) with predicted classifications and $r = .63$ ($n = 118$, $p < .01$) with predicted scores.

Questionnaire Data Corrected for Internal Consistency

More extensive analyses were carried out on questionnaire data corrected for internal consistency. This correction involved only determining whether each subject's response to Question 1, Part III was consistent with his (or her) responses to Questions 4 and 5, Part III (see Appendix A). For example, if a subject indicated he (she) believed he (she) was going to hit 27 targets at record fire (Question 1), but then indicated that the highest number of targets he (she) felt likely to hit was less than 27 (Question 4), the subject's data were not included in the analyses that follow. This also was the case if a subject indicated that he (she) was going to hit 27 targets, but then indicated that the lowest number of targets that he (she) felt likely to hit was greater than 27 (Question 5).

Most of the *r*s computed following this correction were similar in magnitude and direction to the *r*s reported above. Particularly noteworthy, however, was the r between soldiers' predicted and actual record fire scores which showed some improvement, $r = .32$ ($n = 124$, $p < .01$). In addition, the r between reported months since last record fire and record fire performance achieved significance, $r = -.19$ ($n = 123$, $p < .05$). Reported months since last record fire should correlate negatively with record fire performance, at least to the degree that this variable reflects the length of the interval between successive refresher training periods.

Few *r*s declined in strength. Most notable among these were the *r*s between subjects' remembered record fire scores and current record fire scores, $r = .24$ ($n = 97$, $p < .05$), and between soldiers' remembered record fire classifications and current record fire scores, $r = .33$ ($n = 107$, $p < .01$).

As indicated earlier, subjects were asked to predict their shooting classifications (Question 2, Part III). They also were asked to estimate their chances of being correct in these predictions (Question 3, Part III). In general, subjects who estimated their chances of being correct at 90% or 100% were more accurate in their predictions than subjects who gave lower confidence estimates. In particular, predicted shooting classifications and record fire scores correlated $r = -.01$ ($n = 36$, $p > .05$) for subjects who estimated their chances of accurate prediction at less than 60%. This r was .24 ($n = 47$, $.05 < p < .10$) for subjects who estimated their chances of accurate prediction as being between 60% and 80%. And r equalled .42 ($n = 41$, $p < .01$) for subjects who estimated the likelihood of accurate prediction at 90% or 100%. Similarly, the *r*s between predicted and actual record fire scores were nonsignificant for subjects who estimated their chances of accurate prediction at less than 60%, $r = .23$ ($n = 36$, $p > .05$), or between 60% and 80%, $r = .04$ ($n = 47$, $p > .05$). However, this r was highly significant for subjects who estimated their chances of accurate prediction at 90% or 100%, $r = .50$ ($n = 41$, $p < .01$).

Our suspicion was that subjects who were most accurate and who expressed most confidence in their predictions, generally, had more or better information on which to base their predictions than other subjects. In testing this hypothesis, we relied on the data we had compiled on subjects' previous marksmanship training (Part II). We decided not to analyze directly the reported training of subjects who estimated their chances of accurate prediction at 90% or 100%. We felt that this analysis would be misleading. Presumably, subjects who have received little or no training and who knowingly expect to do poorly at record fire may be as accurate and express as much confidence in their predictions as their highly trained counterparts. Instead, we dealt with only the data of subjects who indicated they had (a) not received any M16A1 training since last firing record fire (Question 3, Part II), (b) not had much experience outside the Army firing a rifle or shotgun (Question 4, Part II), and (c) not fired a rifle (other than the M16A1) or a shotgun since last firing record fire (Question 5, Part II). We divided this sample ($n = 57$) into two groups--those who reported firing record fire within the last 12 months ($n = 49$) and those who reported not firing record fire in the last 12 months ($n = 8$). Our assumption was that, among subjects not reporting outside or intervening marksmanship training, those who reported recently firing record fire would have more and better information on which to base their predictions than subjects who reported not firing record fire as recently. Data obtained, generally, were consistent with this assumption. Among those sampled, subjects who reported firing record fire within the last 12 months averaged 46% less error (i.e., absolute, or unsigned error) in their predictions than subjects who reported not firing record fire as recently ($M = 6.82$, $SD = 5.43$ versus $M = 10.63$, $SD = 7.84$), $F(1, 55) = 2.98$, $.05 < p < .10$. No differences were observed between mean confidence estimates in the two groups, although these means were in the expected direction, $F(1, 54) < 1$.

The observation that subjects' confidence estimates were related to the accuracy of their self assessments and that self assessments were related to remembered record fire scores and classifications suggested a multiple regression approach toward predicting record fire performance. Included in the analysis, in addition to subjects' confidence estimates, were predicted shooting classifications and remembered record fire classifications; the dependent variable was record fire scores. The resulting multiple R equalled .34, $F(3, 102) = 4.45$, $p < .01$, and was not substantially better than the simple r between record fire score and remembered record fire classification, $r = .33$. Apparently, confidence ratings cannot be used as a moderator variable in the multivariate prediction of record fire scores, at least not without attempting separate prediction equations for the highest and other self raters.

Questions 4 and 5, Part III, asked subjects to write the "highest" and the "lowest" numbers of targets they felt "likely" to hit at record fire. Only 36.3% ($n = 124$) were correct in their predictions. Even so, 58.1% of the subjects erred by +/- 3 shots or fewer, and 75.9% erred by +/- 7 shots or fewer. Furthermore, these results were documented even though mean ($M = 7.86$; $SD = 6.90$) and median ($Mdn = 6.13$) interval sizes between predicted high and low numbers of targets likely to be hit were reasonably small.

Additionally, each subject was asked if he (she) would be more surprised if he (she) shot "22 or below" or "36 - 40". There was some tendency for better shooters to select "22 or below" and for worse shooters to choose "36 - 40," but the r was only marginally significant, $r = -.16$ ($n = 124$, $.05 < p < .10$). The fact that this question calls for a new response set may account for this marginal r. (Good shooters should have responded "22 or below," while poor shooters should have responded "36 - 40." These responses are the opposite of those that probably were expected, given the context established by previous questions on the questionnaire.) The observation that several subjects later reported being confused by this question is consistent with this interpretation.

Estimation errors (i.e., algebraic, or signed errors) were inversely related to record fire scores, $r = -.73$ ($n = 124$, $p < .01$). An examination of these errors indicated that the predictions of subjects who fired well at record fire, generally, were more accurate than those who fired poorly. This suggests that subjects who performed well also were more skilled at self assessment. Evidence supporting this interpretation comes from Moreland, Miller, and Laucka (1981) who found that students who performed well academically were more accurate in their self assessments than those who performed poorly. On the other hand, this result may merely reflect a general tendency to overestimate shooting ability: 75% of the subjects predicted they would hit more targets than they actually hit. If most subjects thought they would do well, good shooters naturally would show less error in their predictions than poor shooters.

Only five subjects (4%) predicted that they were going to fail to qualify (i.e., hit less than 23 targets). Actually, 34 subjects (27%) failed. More interestingly, however, three of the five subjects who predicted that they were going to fail did in fact fail. And the remaning two who passed, passed by only two points. Conclusions are limited by the small sample size. However, this result suggests that although subjects were biased heavily toward predicting success, those who predicted failure were quite accurate.

DISCUSSION

This research examined the use of self assessments in predicting refresher training requirements. Permanent party soldiers, who reported not firing record fire for 1 to 60 months, first completed a questionnaire designed, in part, to assess their beliefs about how they were going to perform at record fire. These soldiers then fired record fire, after which predicted and actual scores were compared.

Correlations between predicted and actual scores were not high. At best, predictions accounted for only about 10% of the variance associated with record fire scores. Nevertheless, these data must be regarded as encouraging for the following reasons:

Record fire scores are notoriously unreliable. Equipment failure, scorer bias, variations in light conditions, vegetation, and terrain as well as such intangibles as fatigue, motivation, and stress all are believed to contribute to this unreliability. In one experiment, for example, experimenters failed to find a significant correlation between practice record fire scores and record fire scores, $\underline{r} = .10$ ($\underline{n} = 104$). This was true even though these scores were made by the same firers on the same range on successive days (Marcus & Hughes, 1979).

No special instructions were provided subjects as they completed the questionnaire. The idea was to obtain a baseline estimate of soldiers' abilities to assess their own skills under field conditions. Higher correlations probably would have been obtained if special attempts had been made to clarify questions and have subjects check their responses (e.g., Burnside, 1982).

The data obtained appear logically consistent. Predicted and actual performances correlated positively as did predicted and remembered performances. The correlation between predicted and actual scores showed some improvement when data were corrected for internal consistency. Subjects who expressed the most confidence in their predictions were more accurate than other subjects. And, among those sampled, subjects who reported firing record fire within the last 12 months were more accurate in their predictions than subjects who reported not firing record fire as recently.

Future research should focus on improving the accuracy of subjects' self assessments. Several manipulations already are known to improve this accuracy (see Burnside, 1982). They include clarifying the task, conditions for performance, and performance standards (e.g., van Rijn, 1980); designing questions and rating scales so that they are clearly and easily understood (e.g., Burnside, 1982); and training raters before they provide self assessments (e.g., Fischhoff & MacGregor, 1981).

Our experience conducting this research suggests at least four further manipulations:

1. Provide subjects as much information relevant to the formation of accurate self assessments as possible prior to asking them to produce their assessments. Our data suggest, at least, that subjects who have more and better information on which to base their predictions generate more accurate predictions than other subjects. Presumably, this information is conveyed in large measure by descriptions of the task, conditions, and standards. However, it also may be beneficial to remind subjects how they did previously or to refresh their memories for a task using a demonstration.

2. Have subjects provide confidence ratings along with their self assessments. Confidence ratings appear to be a good index of the accuracy of these assessments. In this research, the predicted scores of subjects who estimated their chances of accurate prediction at 90% or 100% accounted for 25% of the variance associated with record fire scores. Similar results have been obtained by Fischhoff and MacGregor (1981), who found that someone is more likely to be right in forecasting the outcome of events when he (she) is "certain" of his (her) forecast than when he (she) is "fairly confident."

3. Check responses for internal consistency. Subjects should be instructed to check their responses for internal consistency as they complete their self-assessment questionnaires. These checks also should precede all other data analyses. Questionnaires that fail these checks either should be returned to subjects for correction or not included in subsequent analyses.

4. Consider asking subjects to make more general self assessments prior to asking them for more detailed responses. The first question we asked subjects was, "How many targets out of 40 do you think you will hit at record fire?" (Question 1, Part III). This may have been a mistake. Subjects might have responded more accurately if they had been led into this estimate, perhaps using questions such as, "What is the lowest number of targets out of 40 you think you might hit today?" or "What is the highest number of targets out of 40 you think you might hit today?" Asking for confidence ratings while continuing to probe for more specific information at least would have enabled us to determine where subjects began to lose confidence in their responses.

Three other results bear some discussion and further empirical testing:

1. One of the best predictors of how well a soldier was going to shoot at record fire was his (her) remembered performance from his (her) most recent record fire. This result suggests that if training remains more or less constant between successive tests or is not sufficient to raise soldiers' retention above some low steady state, then the best predictor of retention may simply be performance on the preceding test. Test scores

usually are readily available to unit commanders. However, with improvements in training, prediction also may improve if both past performance (either self reports or unit records) and self assessments are considered.

2. The vast majority of our subjects were positively biased toward predicting success. Only five subjects predicted failure; however, these subjects were quite accurate in their predictions. Presumably, subjects predicting failure on particular tasks could be identified using a questionnaire. They then could be retrained, even if no one else were considered for retraining at that time.

3. While experience firing a rifle or shotgun did not correlate with performance at record fire, it did correlate with predicted record fire performance. This also was true for those who reported firing a rifle (other than the M16A1) or shotgun since the last record fire and, to a more limited degree, for those reporting some form of M16A1 training since the last record fire. These results support the common assumption that soldiers' confidence in their marksmanship skills increases with experience. This increase in confidence--even if it is false--must be regarded as critical to the survival of the combat infantryman and further underscores the need for excellent Army marksmanship training and refresher training programs.

REFERENCES

Burnside, B. L. (1982, May). Subjective appraisal as a feedback tool (ARI Technical Report 604 , Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Fischhoff, B., & MacGregor, D. (1981, December). Subjective confidence in forecasts (PTR Technical Report PTR-1092-81-12). Woodland Hills, CA: Perceptronics.

Levine, E. L., Flory, A., & Ash, R. A. (1977). Self-assessment in personnel selection. Journal of Applied Psychology, 62, 428-435.

Marcus, A., & Hughes, C. R. (1979, May). An evaluation of a technique for using combat training theater (CTT) for periodic rifle marksmanship proficiency training and qualification (ARI Research Problem Review 79-7). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A075 399)

Moreland, R., Miller, J., & Laucka, F. (1981). Academic achievement and self-evaluation of academic performance. Journal of Education Psychology, 73, 335-344.

Rigg, K. E., & Gray, B. B. (1981, December). The optimization of skill retention through initial training analysis and design (MGA Final Report MGA-2081-MRO-0047). Monterey, CA: McFann-Gray & Associates, Inc.

Rose, A. M., & Ford, P. (1982, April). Acquisition and retention of soldiering skills: Report of year 1 progress (AIR Technical Report AIR/IR-88601-4/82). Washington, DC: American Institutes for Research.

Schendel, J. D., & Hagman, J. D. (1982). On sustaining procedural skills over a prolonged retention interval. Journal of Applied Psychology, 67, 605-610.

Schendel, J. D., Shields, J. L., & Katz, M. S. (1978, September). Retention of motor skills: Review (Technical Paper 313). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A061 338)

van Rijn, P. (1980, June). Self-assessment for personnel examining: An overview (Personnel Research Report 80-14). Washington, DC: U.S. Office of Personnel Management, Personnel and Development Center.

APPENDIX A
SELF-ASSESSMENT QUESTIONNAIRE

DATA REQUIRED BY THE PRIVACY ACT OF 1974

TITLE: Marksmanship Skill Retention and Prediction Questionnaire

PRESCRIBING DIRECTIVE: AR 70-1

AUTHORITY: 10 USC Sec 4503

PURPOSE(S): The data collected with the attached form are to be used for research purposes only.

This is an experimental personnel data collection form developed by the U.S. Army Research Institute for the Behavioral and Social Sciences pursuant to its research mission as prescribed in AR 70-1. When identifiers (name or Social Security Number) are requested they are to be used for administrative and statistical control purposes only. Full confidentiality of the responses will be maintained in the processing of these data.

Your participation in this research is strictly voluntary. Individuals are encouraged to provide complete and accurate information in the interests of the research, but there will be no effect on individuals for not providing all or any part of the information.

PART I: BACKGROUND INFORMATION

1. NAME: ______________________________
   (Last) (First) (MI)

2. SSAN: ______________________________

3. UNIT: ______________________________

PART II: TRAINING

1. Approximately how long has it been since you last fired record fire? _____(Months)

2. How many targets out of 40 did you hit? _______

   If you can't remember, did you fire:

   Expert _____ Sharpshooter _____ Marksman _____ Unqualified _____

3. Have you received any M16A1 training since you last fired record fire? (Check one)

   Yes _____ No _____

   If you checked "No," go to Question 4. If you checked "Yes," check the statement(s) below that best describe the training you received:

   Firing live rounds _____ Dry firing or firing with blanks _____

   Non-firing instruction _____ Weaponeer _____ MILES _____

   Maintenance _____ Other _____

4. Have you had much experience outside the Army firing a rifle or a shotgun (hunting, skeet shooting, and so on)? Answer "No" if you only fired a rifle or a shotgun on a few occasions.

   _____ Yes _____ No

5. Have you fired a rifle (other than the M16A1) or a shotgun since the last time you fired record fire (hunting, skeet shooting, and so on)?

   _____ Yes _____ No

PART III: PREDICTIONS

We want to know what you think you are going to fire at record fire. After record fire we will compare your guess to your actual score to see how good soldiers are at guessing scores. Please answer the following questions as accurately as possible.

1. How many targets out of 40 do you think you will hit at record fire? _____ Of these targets, how many out of 20 do you think you will hit firing from the foxhole? _____ How many out of 20 do you think you will hit firing from the prone? _____

2. Do you think you are going to fire: (Check one)

_____ 36 - 40 (Expert)

_____ 30 - 35 (Sharpshooter)

_____ 23 - 29 (Marksman)

_____ 22 or less (Unqualified)

3. What would you guess your chances are of being correct on the last question? (Check one)

_____ No chance of being correct

_____ 10% chance of being correct

_____ 20% chance of being correct

_____ 30% chance of being correct

_____ 40% chance of being correct

_____ 50% chance of being correct

_____ 60% chance of being correct

_____ 70% chance of being correct

_____ 807 chance of being correct

_____ 90% chance of being correct

_____ 100% chance of being correct

4. What is the highest number of targets out of 40 you feel you are likely to hit today? _____

5. What is the lowest number of targets out of 40 you feel you are likely to hit today?

6. Would you be more surprised if: (Check one)

_____ you shot 22 or below _____ you shot 36 - 40