Technical Report 762

AD-A188 656

Attribute Assessment: Initial Test of Scales for Determining Human Requirements of Military Jobs

Elizabeth P. Smith and Paul G. Rossmeissl

Selection and Classification Technical Area
Manpower and Personnel Research Laboratory

U. S. Army

Research Institute for the Behavioral and Social Sciences

October 1987

Approved for public release; distribution unlimited.

Technical Report 762

Attribute Assessment: Initial Test of Scales for Determining Human Requirements of Military Jobs

Elizabeth P. Smith and Paul G. Rossmeissl

Selection and Classification Technical Area
Lawrence M. Hanser, Chief

Manpower and Personnel Research Laboratory
Newell K. Eaton, Director

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES
5001 Eisenhower Avenue, Alexandria, Virginia 22333-5600

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

October 1987

Army Project Number
2Q263731A792

Manpower and Personnel

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: ARI Technical Report 762	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): ATTRIBUTE ASSESSMENT: INITIAL TEST OF SCALES FOR DETERMINING HUMAN REQUIREMENTS OF MILITARY JOBS	5. TYPE OF REPORT & PERIOD COVERED: Interim Report, January 1984-June 1985
	6. PERFORMING ORG. REPORT NUMBER: --
7. AUTHOR(s): Elizabeth P. Smith and Paul G. Rossmeissl	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 22333-5600	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 2Q263731A792, 2.3.1.H.2
11. CONTROLLING OFFICE NAME AND ADDRESS: U.S. Army Research Institute for the Behavioral and Social Sciences, 5001 Eisenhower Avenue, Alexandria, Virginia 22333-5600	12. REPORT DATE: October 1987
	13. NUMBER OF PAGES: 31
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)

Ability requirements
Job requirements
Ability assessment
Human ability requirements
Job attributes
Attribute requirements

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

The Attribute Assessment Scale (AAS) was developed empirically to enable noncommissioned officers (NCOs) to estimate profiles of human attributes required for different military occupational specialties (MOS). Two experiments were run to test AAS in terms of its interrater agreement and in terms of the differentiation of attributes within and across MOS.

The first experiment included a tri-level performance criterion: NCOs from two MOS provided three ratings of the level required of 22 (Continued)

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

UNCLASSIFIED

 SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ARI Technical Report 762

20. Abstract (Continued)

attributes for work in their own MOS. Results indicated a number of problems, some attributable to the complex criterion (e.g., confusion, response set, ceiling effect).

The second experiment, which included three MOS, corrected for some, but not all of the problems. The conclusion from the two efforts was that interrater agreement and differences observed across MOS, although minimal, were sufficient to warrant further tests of the scale. These future efforts should examine effects of criterion specificity and types of scale authors on the estimates obtained.

ATTRIBUTE ASSESSMENT: INITIAL TEST OF SCALES FOR DETERMINING HUMAN REQUIREMENTS OF MILITARY JOBS

EXECUTIVE SUMMARY

Requirement:

To develop and test an efficient method for linking personal attributes (abilities, interests, characteristics) to job performance as a supplement to empirical validation research.

Procedure:

The Attribute Assessment Scale (AAS), a set of 22 behaviorally anchored rating scales, was developed empirically. Two experiments were conducted to assess the AAS' interrater reliability and ability to differentiate profiles of Skill Level attribute requirements across military occupational specialties (MOS). In Experiment 1, supervisory NCOs from two MOS estimated levels of attributes required for three different performance levels. In Experiment 2, NCOs from three MOS estimated requirements for average performance only. These NCOs also participated in lengthy discussions about the method. Analyses of variance were performed to determine overall reliability coefficients and MOS differences. Additional reliability for coefficients for individual attributes were calculated in Experiment 2.

Findings:

Interrater reliabilities obtained in Experiment 1 were disappointing. Although coefficients were moderate, estimates of the reliability that would be obtained if smaller, more efficient numbers of raters were used were unacceptably low. On the other hand, coefficients obtained with small samples in Experiment 2 were quite good. Significant MOS differences in profiles of attributes were not found in either case. Several potential sources for this lack of differentiation were identified.

Utilization of Findings:

Results of this research will be used to modify the research design to enable additional testing of the Attribute Assessment Scale and subsequent development of an alternative approach if necessary.

FOREWORD

The Army Research Institute is currently engaged in an extensive long-term research effort, Project A, to improve the selection, classification, and utilization of enlisted personnel. Empirical validity investigations such as this, however, are costly and time consuming. To supplement them, we need other methods for optimizing people-to-job matches, including the use of empirical findings in new or additional ways. This report describes exploratory research to evaluate the reliability of one adjunct method. The Attribute Assessment Scale, a set of 22 rating scales, was developed for supervisors to estimate the levels of human attributes, i.e., abilities, interests, and characteristics, required to do different military jobs. It was designed so that the attributes corresponded to the constructs being considered as predictor measures by Project A, to enable matching job-requirements profiles to profiles of predictor test scores. These early findings showed that, although the instrument may have moderate to good agreement among raters (i.e., interrater reliability), the resulting attribute profiles do not differentiate military occupational specialties (MOS). These limited results suggest that additional research with modifications should be conducted.

EDGAR M. JOHNSON
Technical Director

ATTRIBUTE ASSESSMENT: INITIAL TEST OF SCALES FOR DETERMINING HUMAN REQUIREMENTS OF MILITARY JOBS

CONTENTS

LIST OF TABLES

ATTRIBUTE ASSESSMENT: INITIAL TEST OF SCALES FOR DETERMINING HUMAN REQUIREMENTS OF MILITARY JOBS

INTRODUCTION

Conducting empirical validity investigations to predict job performance is not always feasible. Even when empirical approaches are undertaken, such as the ongoing Army Research Institute's (ARI) Project A to improve the selection, classification, and utilization of enlisted personnel, it is rarely possible to include all jobs within an organization. Given the complexities of empirical validation, it is necessary to develop other methods for matching people to jobs and optimizing their performance.

One approach is to obtain rational estimates of the human attributes (i.e., abilities, characteristics, and interests) which are required for successful job performance. When gathered systematically from qualified judges, these estimates can be summarized as profiles of required attributes. Then, measures of individuals' attributes can be matched to such profiles for selection and classification purposes. In addition, knowledge of required attributes is potentially useful for (a) designing new systems and training programs that are within the capacities of available personnel and (b) generalizing empirical validity data to new and different jobs, by grouping them on the basis of similarity of attribute profiles (Fleishman, 1982; Pearlman, 1980). The latter application is especially pertinent to Project A, which is collecting validity data for only 19 Military Occupational Specialties (MOS).

One method of determining ability requirements is the rating scale approach developed by Fleishman and his associates based on a taxonomy of 40 cognitive, perceptual, physical and psychomotor abilities. A comprehensive summary of this work included in Fleishman and Quaintance (1984) indicates that the scales have construct validity and yield reliable estimates of ability requirements. With these scales, a rater decides if an ability is necessary for errorless job performance, and, if so, estimates the level required on a 7-point, behaviorally anchored scale.

Recent efforts (Rossmeissl & Dohme, 1982) to test the use of rating scale assessment of attribute requirements for Army MOS followed directly from Fleishman's work. The researchers viewed the universal anchor points of Fleishman's scales as too general for application to Army MOS. With the assistance of incumbents in the field, they developed new scales for his taxonomy which had anchor statements directly relevant to specific Army jobs in aviation. They obtained high interrater reliability and reliable discrimination among attributes. Ratings did not discriminate across the four targets tasks (i.e., different helicopter missions), however, raising the question of how variable tasks or jobs must be to observe differences in necessary attributes.

The positive results from the aviation research led to the development of a computerized system for rating aptitude requirements, the Job Assessment Software System (JASS) (Rossmeissl, Tillman, Rigg, & Best, 1984). This system was built upon the decision flow diagram method developed by Mallamad, Levine, and Fleishman (1980). With this method, raters make a series of simple Yes-No responses to questions about abilities necessary for performance of the target job or task. Specific response patterns identify only those abilities which are required. Rating cales are provided for these abilities only.

An initial test of the JASS method (Rossmeissl, Kostyla, & Tillman, 1983) compared JASS procedures to paper-and-pencil scales (without flow diagrams) on what abilities were selected as necessary and on interrater reliabilities. Overall, the results indicated that the computer method was as reliable as the paper-and-pencil method. That is, both types of scales yielded high interrater reliabilities and were able to discriminate reliably among attributes. A subsequent test of the computerized rating scales (Olson & Hanser, 1983) with Army Infantry MOS also provided favorable results: interrater reliability within groups was moderate and different profiles were obtained across the four MOS examined. These preliminary results suggest that further investigation into rating scales to assess attribute requirements of Army jobs, especially with a view toward computerization, has merit.

Early outcomes from Project A provided an opportunity to develop a set of rating scales based on a new taxonomy of human attributes. An expert judgment task (Wing, Peterson, & Hoffman, 1984) obtained estimates of validity for 53 predictors against 72 criterion constructs from 35 personnel psychologists. Factor analysis of the data yielded 21 clusters of the 53 cognitive, perceptual, psychomotor, temperament, and interest predictor variables. A predictor test battery based on these 21 clusters has been developed and is being validated. The purpose of this paper is to discuss the initial construction and testing of a new set of scales for estimating job requirements which is based on these 21 clusters (hereafter called "attributes"). As more data become available, it is expected that the taxonomy of predictors (and test battery) may change. The rating scales will be revised to reflect these changes.

A set of scales based on the Project A taxonomy has several potential advantages over the Fleishman ones. The most salient feature is that obtained profiles of attribute requirements will directly correspond to Project A validity data. It will include temperament and interest measures that are not among the Fleishman scales and will not include those attributes/abilities for which no predictor tests are given. Thus extraneous data collection can be avoided. The new set of scales was designed to be used by work supervisors rather than personnel psychologists and contains primarily Army-specific behavioral anchors with only about half as many attributes to rate as Fleishman's.

Research Problem

For any rating scales to be useful in practice, they must give reliable and valid scores. This report describes two experiments designed to examine issues related to the reliability of the ratings from the new set of scales which collectively form the Attribute Assessment Scale (AAS). Ratings were obtained from supervisors as Subject Matter Experts (SMEs). Validity investigations will occur later, when acceptable reliability has been established. The first experiment considered the following questions.

First, how closely do raters agree, i.e., how high is interrater reliability? From the AAS, we obtain a set of mean ratings over all raters (SMEs) which serve as estimates of the actual requirements of an MOS. If the individual responses vary widely, then means based on these responses are likely to vary greatly from SME sample to sample. Thus, such estimates would be of little or no value.

Second, how well do the scales differentiate across attributes within a job and across the attribute profiles of different jobs? All attributes will not be required at the same level within a job. Thus, estimates obtained from ratings on the AAS should not yield flat profiles. More importantly, different MOS should vary in their patterns of required attributes. To be of use, AAS must produce dissimilar profiles for different MOS.

The final question addressed in Experiment 1 was: Can the scales be used to identify attributes for which differences in level of the attribute most influence performance? For some attributes, higher levels may be required for better performance. For others, once a minimal requirment is met, having a greater amount of the attribute may have no additional effect on performance. Thus, it would be beneficial to be able to determine those attributes for which variability in the attribute within individuals would produce the most variability in performance. Also, it was hoped that by varying the performance level, data could be obtained to help determine which performance criteria should be used in future revisions of the instrument.

Several problems were uncovered during the first experiment. The second experiment attempted to address issues primarily related to the criterion, especially its tri-level nature. Only one performance level was rated and a written job description was provided to the SME's. Thus Experiment II focused on the first two concerns above: interrater reliability and differentiation of attribute ratings within and across MOS. This experiment also looked at interrater reliability within each attribute as well as across all scales.

DEVELOPMENT OF THE ATTRIBUTE ASSESSMENT SCALE

The Attribute Assessment Scale (AAS), consists of a set of 22 behaviorally-anchored scales. Scales were created for 20 of the 21 attributes in the Project A taxonomy plus two additional attributes, Stamina and Physical Strength. The latter were thought to enhance face validity. A scale for Enterprising Interests, the remaining Project A attribute, was eliminated from the instruments. It was impossible to generate items for this attribute which were sufficiently different from those falling under Self-Esteem/Leadership to enable SMEs to distinguish the two attributes. The names of the attributes were modified from the original Wing, et. al. (1984) cluster labeling for better comprehension by SMEs. The attributes included in the AAS are given in Table 1. Their definitions are provided in Appendix A.

To construct the scales, comprehensive definitions for the attributes were developed so as to be readily understandable by people who were not trained in personnel research. A pool of items for potential anchors (i.e., behavioral statements) was generated. Ten items per attribute were ultimately selected, after screening by two to four other researchers. These items were presented with the appropriate definition in an anchor-rating instrument. Initially, 26 NCOs from either the Administrative Specialist (71L) or Military Police (95B) MOS rated each item on the amount of the attribute represented by or needed for the behavior described. Items with mean ratings that were the highest, lowest, and closest to 4.0 (midpoint) that also had a standard deviation less than 1.5 were selected as scale anchors. Using these criteria, scales could be created for only 11 attributes.

After identifying difficulties related to (a) task comprehension, (b) response format, (c) failure of raters to differentiate effectively among items, and (d) a few of the definitions and items themselves, we revised the anchor-rating instrument and administration procedures, adding a 15-minute training period. This instrument was given to another sample of NCOs (N=28) from the same two MOS. From the second administration, using the criteria indicated above, three anchors were obtained for all but two of the attributes (Social Interaction and Stress Reaction for which only two anchors were selected) to form the AAS. Sample pages from the anchor rating instruments for both samples, as from well as the AAS, are included in the Appendices B-D.

Table 1

Attributes Included in the Attribute Assessment Scale

Cognitive/perceptual	Physical/Psychomotor
Verbal Ability	Physical Strength
Memory	Stamina
Reasoning Ability	Multilimb Coordination
Number Facility	Dexterity
Mechanical Comprehension	Steadiness/Precision
Information Processing	
Closure	**Noncognitive**
Visualization	
Perceptual Speed & Accuracy	Social Interaction
	Stress Tolerance
	Conscientiousness
	Work Orientation
	Self Esteem/Leadership
	Athletic Ability/Energy
	Realistic Interests
	Investigatve Interests

IMPLEMENTATION OF THE ATTRIBUTE ASSESSMENT SCALE: EXPERIMENT 1

Method

Subjects. Thirty-six Non-commissioned Officers (NCOs) from the Cannon Crewman (13B) MOS and 39 NCOs from the Motor Transport Operator (64C) MOS, all males located overseas, participated as Subject Matter Experts (SMEs).

Instrument. This research used the AAS described above. The instrument has one page per attribute, with the definition at the top. For this experiment, there were three 7-point vertical scales, placed side-by-side, to enable three responses. A zero-point was added to indicate the attribute was not required at all. SMEs circled the number corresponding to the appropriate level needed for their job.

Procedure. SMEs rated the level of each of the 22 attributes that is required to perform Skill Level 1 (entry level) work under combat-readiness conditions in their own MOS for three performance levels: at the 15th, 50th, and 85th percentiles. In addition to the written instructions, SMEs received extensive training in how to complete the task. This included a step-by-step demonstration of the actual rating process using the anchors as guides. Training and responses to questions took about an hour. Early ratings were checked to ensure comprehension of the directions before raters proceeded with the rest of the task. Ratings took about 30-45 minutes. Comments were solicited during a short debriefing period.

Analyses. We first calculated means and standard deviations of the ratings by MOS. Intraclass correlation coefficients (ICCs) were calculated from Raters X Attributes ANOVAs over all attributes and separately for the three major domains (i.e., cognitive/perceptual, physical/psychomotor, and noncognitive) for each of the three performance levels. One form of ICC estimates the reliability of the mean ratings ($\underline{r}_k$; k = number of raters). This was calculated by $(MS_{att} - MS_{err})/MS_{att}$. A second form estimates the reliability of a single rating ($\underline{r}_1$) and provides an index of interrater reliability. The formula for $\underline{r}_k$, is $(MS_{att} - MS_{err})/MS_{att} - (k-1)\ MS_{err}$. Finally, we performed an MOS X Attributes X Performance Levels univariate repeated-measures ANOVA to test for Attribute and MOS profile differences.

Results

Eight Motor Transport Operators were eliminated from the analyses due to the logical inconsistency of their data. Table 2 contains means and standard deviations of the ratings. The $\underline{r}_k$ coefficients over all attributes were, in increasing order by performance level, .75, .77, and .69 for Cannon Crewmen (k=36) and .74, .74, and .69 for Motor Transport Operators (k=31). For the major domains, (cognitive/perceptual, physical psychomotor, and non cognitive $\underline{r}_k$ coefficients ranged from .61 to .79 across performance levels and MOS. There were two exceptions to this: Physical/ psychomotor reliabilities were very low for both MOS at the 85th percentile [$\underline{r}_{36}$=.13; r_{31}=.38] performance level. The $\underline{r}_1$ coefficients were extremely small. Table 3 contains all $\underline{r}_k$ and $\underline{r}_1$ coefficients.

Table 2

Means and Standard Deviations of Attribute Requirements for Cannon Crewman and Motor Vehicle Operator MOS at Three Performance Levels for Experiment 1

Attributes	MOS[a]	Performance Level					
		15th Percentile		50th Percentile		85th Percentile	
Cognitive/Perceptual		M	SD	M	SD	M	SD
Verbal Ability	C	2.86	(.93)	4.17	(.91)	5.33	(1.10)
	D	2.87	(.83)	4.32	(.85)	5.32	(1.08)
Memory	C	2.44	(1.18)	3.89	(.89)	5.53	(1.11)
	D	3.26	(1.26)	4.26	(1.03)	5.16	(1.27)
Reasoning Ability	C	2.78	(1.31)	3.94	(1.17)	4.89	(1.33)
	D	2.45	(.96)	3.77	(1.02)	5.00	(1.32)
Number Facility	C	2.06	(1.12)	3.47	(1.16)	5.08	(1.52)
	D	2.48	(1.18)	3.90	(1.30)	4.94	(1.48)
Mechanical Comprehension	C	2.86	(1.36)	4.39	(1.18)	5.50	(1.08)
	D	2.97	(1.33)	4.23	(1.14)	5.20	(1.00)
Information Processing	C	2.50	(1.08)	3.58	(1.30)	4.81	(1.37)
	D	2.68	(1.30)	3.90	(1.08)	5.03	(1.17)
Closure	C	2.78	(1.33)	4.08	(1.25)	4.89	(1.41)
	D	2.68	(1.42)	3.87	(1.45)	4.61	(1.67)
Visualization	C	2.33	(1.15)	3.58	(1.30)	4.69	(1.51)
	D	2.29	(1.30)	3.42	(1.43)	4.26	(1.84)
Perceptual Speed & Accuracy	C	2.75	(1.52)	3.97	(1.38)	4.94	(1.31)
	D	2.97	(1.33)	4.16	(1.29)	4.71	(1.40)
Physical/Psychomotor							
Physical Strength	C	3.75	(1.32)	4.89	(1.14)	5.67	(1.17)
	D	3.58	(1.39)	4.61	(1.20)	5.32	(1.25)
Stamina	C	3.06	(1.45)	4.53	(1.11)	5.69	(1.19)
	D	2.68	(1.30)	3.94	(1.41)	4.84	(1.63)
Multilimb Coordination	C	2.80	(1.47)	4.20	(1.21)	5.26	(1.40)
	D	3.34	(1.54)	4.45	(1.24)	5.48	(1.12)
Dextery	C	3.00	(1.35)	4.47	(1.08)	5.50	(1.08)
	D	3.10	(1.45)	4.19	(1.08)	5.00	(1.39)
Steadiness/Precision	C	2.83	(1.40)	4.08	(1.25)	5.47	(1.36)
	D	3.06	(1.29)	4.52	(1.09)	5.29	(1.07)
Noncognitive							
Social Interaction	C	3.14	(1.62)	4.44	(1.59)	5.34	(1.70)
	D	2.58	(1.71)	3.74	(1.44)	4.65	(1.70)
Stress Tolerance	C	3.03	(1.50)	4.22	(1.27)	5.12	(1.43)
	D	3.10	(1.47)	4.27	(1.34)	5.27	(1.20)
Conscientiousness	C	2.66	(1.24)	4.09	(.89)	5.31	(.99)
	D	3.19	(1.47)	4.35	(1.11)	4.97	(1.25)
Work Orientation	C	2.91	(1.46)	4.29	(1.18)	5.54	(1.17)
	D	2.90	(1.45)	4.16	(1.10)	5.39	(1.17)
Self Esteem/Leadership	C	3.00	(1.26)	4.25	(1.20)	5.47	(1.21)
	D	2.48	(1.55)	3.84	(1.37)	5.00	(1.41)
Athletic Ability/Energy	C	2.89	(1.35)	3.92	(1.16)	4.94	(1.19)
	D	2.87	(1.55)	3.73	(1.48)	4.33	(1.71)
Realistic Interests	C	2.54	(1.40)	3.71	(1.25)	4.94	(1.66)
	D	2.48	(1.12)	3.61	(.88)	4.61	(.99)
Investigative Interests	C	2.00	(1.43)	3.44	(1.61)	4.61	(1.78)
	D	1.97	(1.40)	3.26	(1.50)	4.16	(1.93)

[a] C = Cannon Crewman
D = Motor Vehicle Operator (Driver)

Table 3

Intraclass Correlation Coefficients (ICCs) for Mean and Individual Ratings of Attribute Requirements by MOS and Performance Levels for Experiment 1

MOS		Performance Level					
		15th %tile		50th %tile		85th %tile	
		13B[a]	64C	13B	64C	13B	64C
All Attributes	r_k	.75	.74	.77	.74	.69	.68
	r_1	.08	.09	.09	.08	.06	.07
Cognitive/ Perceptual	r_k	.61	.64	.71	.68	.71	.64
	r_1	.04	.05	.06	.07	.06	.05
Psychomotor/ Physical	r_k	.79	.73	.75	.64	.14	.38
	r_1	.10	.08	.08	.05	.00	.02
Noncognitive	r_k	.73	.78	.70	.76	.62	.75
	r_1	.07	.10	.06	.09	.03	.09

[a] 13B = Cannon Crewman (k = 36)
64C = Truck Driver (k = 31)

None of the effects involving MOS for the MOS X Attributes X Performance Levels ANOVA were significant. There were significant main effects for Attributes [F (21,1365) = 6.98; p = .0000] and Performance Levels [F (2,130) = 398.36; p = .0000] and a significant effect for the Attributes X Performance Levels interaction [F (42,2730) = 2.51; p = .0000]. Scheffes' comparisons between means within performance levels by MOS indicated significant differences between only the highest and lowest means, which ranged from 1.75 to 1.09.

Discussion

In comparison to the very high Intraclass Correlation Coefficients (ICCs) obtained by Fleishman and associates and by Rossmeissl and his associates, the ICCs from this research are weak, especially since around 30 raters are needed to obtain coefficients of at least .60. ICCs are based on variance components. As such, low (or uninterpretable) reliabilities result if there is too great a between-subjects variance and/or too little within-subjects variance. The low reliabilities obtained here appear to be a function of both.

Previous research on ability assessment has found mean ratings that varied from very low (even "Not required") to very high (7) across attributes.

This was not the case here. The inclusion of three performance levels may have had a strong, negative impact on these particular results. The demands of the task appeared to impose a unique kind of restriction in the range of possible ratings. That is, the effective range of ratings within levels covered only two or three points rather than the entire seven points. This outcome served to reduce within-subjects variability, as all ratings fell close together. Although SMEs were clearly advised not to respond according to belief that "better must mean more", the mean ratings suggest that a demand characteristic was created by the instructions to rate at three levels. The result was ratings of attribute levels which correspond to level of performance, with ceiling effects occurring at the highest level. These effects would explain the extremely low reliabilities for Physical/Psychomotor attributes at the 85th percentile.

The fact that attribute requirements were elicited for three performance levels also may have clouded the findings in another way and reduced interrater agreement, i.e., increased between-subjects variance. Although definitions were provided for the three performance levels, how the SMEs actually interpreted these definitions was unknown. SMEs may have had different interpretations of the attributes from our definitions as well as from one another. For example, their verbal reports seemed to indicate some tendency to interpret performance levels in terms of particular soldiers in their charge, rather than from a more general (and shared) view of job performance at a particular level. It is also possible that they tended to rate attributes in terms of the characteristics of someone who performed at that level, rather than in terms of the actual requirements of the job. The performance criterion, then, was more ambiguous than expected, pointing out a clear need for a very specific definition of the criterion. It was apparent that understanding the task requirements -- what was meant by the performance levels and how to do three ratings at a time -- took more time and energy than actually doing the ratings. In short, the use of three performance levels may have made the task harder than was intended, and interfered with the SMEs' ability to rate true requirements.

Two other factors may have contributed to low interrater agreement. SMEs were not given written descriptions of what they were to rate. Instead they were asked to decide individually the nature and content of entry level work and, specifically, what it required in terms of attributes. Moreover, they were to rate the whole job -- all work within all duty positions -- and not just some specific task or set of tasks. This very broad scope allowed considerable opportunity for variance. As a result of personal experiences and/or selective memory, the SMEs could differ a great deal in what they were evaluating. Higher interrater agreement might be expected for narrower areas of consideration. In addition, some SMEs found the scale anchors frustrating rather than helpful. Raters appeared to have difficulty using anchors as reference points for comparing tasks within their MOS. Some tended to evaluate the job in terms of whether the exact tasks depicted were or were not an actual part of the job. With anchors that depicted common soldier tasks, some SMEs had problems separating the overall soldier requirements from the specific job requirements. Thus, although very familiar behaviors were thought to be the best for illustrating a level of an attribute, this was not necessarily the case.

The results of the ANOVA indicate that attribute profiles for the two MOS are not significantly different. While it was expected that differences among Attributes, Performance levels, and their interactions would be obtained, the fact that only the highest and lowest mean comparisons were significant attest to the general lack of discrimination among the ratings.

Despite these problems, the data provide some useful information. The minimal differences which do occur suggest that some differences (as well as similarities) between MOS may exist, but may be masked in the present research for the reasons previously noted. In addition, rank orders of the magnitude of ratings were different for both MOS at all performance levels, again suggesting there may be some differences in patterns of attributes which need further examination. For instance, at the 85th percentile, Verbal Ability ranked tenth for Cannon Crewman but third for Motor Vehicle Operator, while Stamina ranked first and fifteenth respectively. That is, the five attributes with the highest ratings, are different for each MOS. It is important to note, however, that the top five attributes are not necessarily the most important attributes: They are ranked on level of required attribute only and not on relative importance of the attribute.

In summary, NCOs appeared to understand in general how to use the set of scales to rate job requirements. The requirement to produce three sets of ratings simultaneously, however, could have created some problems. The actual physical arrangement of the scales on the page confused people. Also, it seemed to impose limits on the magnitude of ratings assigned. Given the expanse of the criterion to be rated -- the entire MOS at Skill Level 1 -- and the limitations created by the design itself -- different performance levels -- the obtained indices of interrater agreement are reasonable.

These findings suggested that better reliability estimates might be obtained with fewer raters if SMEs were asked to rate requirements for a single performance level; i.e., to estimate the minimum level of an attribute required to perform the job successfully. Elimination of the restriction in range of ratings which was created by including three performance levels, should yield better discrimination among the attributes within MOS, and differences in attribute profiles across MOS. Further, we thought that by focusing raters' attention on evaluating a specific task, a well-defined set of tasks, or a written job description would yield better reliability.

Finally, results suggested that more reliable ratings might be obtained by changing to a generic set of scale anchors (e.g., very low, low, moderate, etc.) or otherwise replacing the present behavorial anchors. Experiment 2 provided us with the opportunity to try out changes regarding performance level and job description with a very small sample.

IMPLEMENTATION OF THE ATTRIBUTE ASSESSMENT: EXPERIMENT 2

Method

Sample. SMEs, all male, were 3 officers and 5 NCOs from the Ammunition Specialist (55B) MOS, 4 NCOs from the Motor Vehicle Transport (64C) MOS, and 6 officers and 3 NCOs from Administrative Specialist (71L) MOS.

Instrument. The AAS was the same as in Experiment 1, but with only one vertical scale. SMEs gave a single rating of the level of each attribute required for "average" performance.

Procedure. We met with SMEs in small groups separately by MOS and status (commissioned vs. noncommissioned officers) for a 2-hour session. In the first hour, after a brief explanation and training period, they rated the attribute requirements for entry level work (i.e., Skill Level 1) in their MOS. During the second hour, we discussed any problems that they had in completing the task and specific issues related to interpretation of "average" performance, confidence in their responses, and ways to improve the procedures. Finally, we derived group consensus ratings of requirements.

Analyses. We first calculated means and standard deviations of the ratings by MOS. Second, we calculated within-group interrater reliability coefficients (r_{wg}s) (James, Demaree, & Wolf, 1984) to determine the interrater agreement within each attribute. R_{wg} is a function of observed variance and variance that would be expected if ratings were due solely to random errors of measurement. Unlike ICCs, r_{wg} is not negatively influenced by too little between-subjects variance. That is, perfect agreement among raters would yield poor ICCs. Next, we calculated ICCs (both r_k and r_1 from Raters X Attributes ANOVAs for each MOS. Finally, we ran an MOS X Attributes univariate repeated measures ANOVA to look at differences within and across MOS profiles. We did not calculate ICCs for the three domains since we calculated r_{wg}s for individual attributes.

Results

Means and standard deviations of the ratings are presented in Table 4. Twenty-two within-group reliability estimates (r_{wg}s) were calculated separately for each MOS. These coefficients are presented in Table 5. Note that Motor Vehicle Transport Operator SMEs (n=4) attained the best set of coefficients, ranging from .11 to 1.0 (M=.71). Only one r_{wg} is less than .40, and 17 (77%) are greater than or equal to .60. For Administrative Specialist SMEs (n=9), r_{wg}s ranged from .27 to .94 (M=66), with 12 (55%) greater than or equal to .60. There is less agreement among the Ammunition Specialist SMEs (n=8). The r_{wg}s ranged from -.05 to .93 (M=58), with two negative values (Steadiness and Social Interaction), which are not interpretable as reliability coefficients. Thirteen (59%) are greater than or equal to .60. As can be seen in the table, r_{wg}s vary greatly across the three MOS with no observable pattern; e.g., the three r_{wg}s for an attribute are not consistently high (low) for all MOS.

Table 4

Means and Standard Deviations of Attribute Requirements for Experiment 2

Attributes	Ammunition Specialist n = 8		Motor Transport Operator n = 4		Administrative Specialist n = 9	
Cognitive/Perceptual						
Verbal Ability	3.25	(1.04)	3.75	(1.50)	4.22	(0.67)
Memory	3.63	(1.19)	4.50	(0.58)	4.00	(0.87)
Reasoning	3.50	(1.20)	3.50	(1.00)	3.11	(1.36)
Number Facility	3.75	(1.58)	3.00	(1.41)	3.44	(1.13)
Mechanical Comprehension	3.50	(1.31)	4.75	(0.50)	2.78	(1.72)
Information Processing	3.38	(0.74)	4.25	(0.96)	3.78	(0.83)
Closure	3.63	(0.52)	4.25	(1.26)	3.22	(1.48)
Visualization	3.38	(0.92)	4.00	(0.82)	2.89	(1.05)
Perceptual Speed & Accuracy	4.50	(0.53)	4.50	(1.00)	5.33	(0.87)
Physical/Psychomotor						
Physical Strength	4.50	(0.93)	5.50	(1.00)	3.89	(1.36)
Stamina	4.50	(1.07)	5.25	(0.96)	4.33	(1.32)
Multilimb Coordination	4.00	(1.31)	5.00	(0.00)	4.00	(0.50)
Dexterity	4.38	(1.19)	5.00	(0.82)	4.56	(1.01)
Steadiness/Precision	4.13	(2.10)	4.75	(1.26)	4.89	(1.36)
Noncognitive						
Social Interaction	4.25	(2.05)	5.50	(1.29)	4.67	(1.41)
Stress Tolerance	4.00	(0.93)	5.25	(0.96)	4.33	(1.00)
Conscientiousness	4.75	(1.91)	5.50	(1.00)	4.67	(0.71)
Work Orientation	5.13	(1.46)	6.50	(0.58)	5.11	(1.45)
Self Esteem/Leadership	4.38	(1.60)	5.75	(0.50)	4.67	(1.00)
Athletic Ability/Energy	3.75	(1.75)	4.00	(1.15)	4.44	(1.51)
Realistic Interests	3.63	(0.74)	3.75	(1.89)	3.33	(1.12)
Investigative Interests	3.00	(0.76)	2.25	(1.50)	2.78	(0.67)

Intraclass correlation coefficents across all attributes for indivdual rater (r_1) calculated from three Raters X Attributes ANOVAs are of the same order of magnitude as the r_{wg}s. R_1 coefficients for Motor Vehicle Transport Operator, Administrative Specialist, and Ammunition Specialist SMEs are, in order, .40, .36, .10. Estimates of the interrater reliability of the mean ratings (r_k) are, in the same order, .73, .84, and .43.

Table 5

Within-group Reliability Coefficients for Attributes Requirements for Experiment 2

Attributes	Ammunition Specialist n=8	Motor Transport Operator n=4	Administrative Specialist n=9
Cognitive/Perceptual			
Verbal Ability	.73	.44	.89
Memory	.65	.92	.81
Reasoning	.65	.75	.54
Number Facility	.38	.50	.32
Mechanical Comprehension	.57	.94	.27
Information Processing	.86	.77	.83
Closure	.93	.61	.45
Visualization	.79	.84	.72
Perceptual Speed & Accuracy	.93	.75	.81
Physical/Psychomotor			
Physical Strength	.79	.75	.54
Stamina	.72	.77	.56
Multilimb Coordination	.57	1.00	.94
Dexterity	.65	.83	.74
Steadiness/Precision	-.10	.61	.54
Noncognitive			
Social Interaction	-.05	.58	.50
Stress Tolerance	.79	.77	.75
Conscientiousness	.09	.75	.88
Work Orientation	.47	.92	.47
Self Esteem/Leadership	.36	.94	.75
Athletic Ability/Energy	.23	.67	.54
Realistic Interests	.86	.11	.69
Investigative Interests	.86	.44	.89
	$\bar{r}_{wg}$=.58	$\bar{r}_{wg}$=.71	$\bar{r}_{wg}$=.66

Discussion

The interrater reliabilities obtained in Experiment 2 were far more acceptable than in Experiment 1. Even with small samples, the majority of the reliability coefficients were greater than .60. Motor Transport Operator SMEs (n=4) attained the highest coefficients. This finding may be due partly to the fact that the job is more similar across assignments than others. Administrative Specialist SMEs were quick to point out how very different the work was within just the few assignments they represented. For example, some SMEs supervised jobs which required considerable people contact (i.e., "customer service") while others supervised soldiers who worked in isolation. The high interrater agreement among the Motor Transport Operator SMEs may also be due, however, to the fact that the small sample size limited the amount of possible variability in ratings. One would expect that similarity of the actual job across assignments would also be the case for Ammunition Specialists. These SMEs, on the other hand, had the lowest reliability coefficients. One explanation for this result is that the conditions under which these SMEs completed the rating task were less than ideal: Several of them arrived at odd times and completed the task while other groups were in session. Thus, they may not have given the same effort or attention to the task as other subjects.

ANOVA results indicated no significant differences in profiles across MOS. In interpreting this, first, it must be remembered that we were dealing with exceptionally small n's. Second, post-rating discussions indicated that use of criterion of "average" performance may have confounded the results. That is, it was not a good choice of terms. There was some tendency of the ratings to converge on the midpoint or "average" of the 7-point scale since SMEs confused average performance with average level requirements. SMEs also indicated they would not be satisfied if the majority of the new recruits had profiles equivalent to their rated requirements: They would want people with higher levels of certain attributes than they had given as required for average performance. In some sense, they seemed to be describing what they consider to be the average soldier, who is not necessarily performing very successfully on the job.

The discussions also revealed that many SMEs thought that generic (or even civilian) anchors might be more effective than the Army specific anchors we had. Although some SMEs thought them helpful, others found the anchors more of a hindrance or distraction. This was especially true of those that reflected common soldier tasks. In addition, as stated above, we found that just within the small groups of SMEs we had, the actual jobs they supervised were very different and so emphasized different attributes. Thus, the job description from Army Regulation 611-201 provided only minimal help. SMEs thought that ratings of requirements for specific component tasks would lead to greater consensus among raters.

CONCLUSIONS

This research illustrates a number of problems in using a rating scale approach for estimating the attributes or abilities required by Army jobs. One of these problems is determining the appropriate level of job performance to be used as the basis for the attribute estimates. We have shown that both multiple performance levels or the term "average performance" can be confusing to Army SMEs. On the other hand, the criteria used by Fleishman, namely that of error-free performance, seems to be extreme and would probably lead to excessively high attribute requirement estimates. Additional research is needed to arrive at a base performance level that is readily understood by Army SMEs and makes sense in the context of personnel selection and classification.

Another problem concerns the nature of the anchor statements on the rating scales. We had thought that the anchors that were specific to the Army would be easier to use than a more generic form of anchor. Our present experience did not support this expectation. Many SMEs found the anchors, particularly those based upon Army common tasks to be confusing. Examples of such confusion included SMEs rating the anchors rather than the job and failing to focus on the unique elements of MOS performance when making their ratings. In this case perhaps generic (non-Army) anchors or no anchors at all would be the best approach.

Further research is also needed to determine just what should be rated by the SMEs. The present research indicated that the requirement of rating the whole job or MOS may not be appropriate. The SMEs stated that a soldiers duties within an MOS could vary considerably as a function of where the soldier was assigned and which particular duty position he or she was holding. For example, a soldier in MOS 71L (Administrative Specialist) could be a member of an office pool whose sole duty is typing, or be alone in in a office and be totally responsible for all of its activities. Perhaps one solution to this dilemma is to determine the key aspects or essential tasks for each MOS and have the SMEs rate the attributes required performance on those tasks. In this manner one could assume that all of the ratings were being made against a common metric.

A major concern with the rating method is the uniformity of attribute profiles across MOS. In both experiments there was little or no differences in the attribute requirements from one MOS to another. A major goal of this research effort was to uncover a method for determining differences among MOS so that the Army applicants could be classified into appropriate MOS. Unless the method can show differences among MOS it is useless in this regard. It is possible that if the procedural problems noted above are remedied, the method may differentiate among the MOS attribute requirements. But it is also possible that the duties among Army MOS are so similar that a rating scale approach is not sensitive enough to capture the differences in the attributes required for successful performance.

REFERENCES

Fleishman, E. A. (1982). Systems for describing human tasks. American Psychologist, 37, 821-834.

Fleishman, E. A. & Quaintance, M. K (1984). Taxonomies of human perform ance. Orlando, FL: Academic Press Inc.

James, L. R., Demaree, R. G, & Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. Journal of Applied Psychology, 69, 85-98.

Mallamad, S. M., Levine, J. M., & Fleishman, E. A. (1980). Identifying ability requirements by decision flow diagrams. Human Factors, 22, 57-68.

Olson, D. M., & Hanser, L. M. (1983, October). Examination of ability requirements for the Infantry Career Management Field. Proceedings of the 25th Annual Conference of the Military Testing Association, Gulf Shores, AL.

Pearlman, K. (1980). Job families: A review and discussion of their implications for personnel selection. Psychological Bulletin, 87, 1-28.

Rossmeissl, P. G., Dohme, J. A. (1982). Using rating scales to determine the aptitude requirements of Army systems. Proceedings of the 24th Annual Conference of the Military Testing Association, San Antonio, TX.

Rossmeissl, P. G., & Kostyla, S. J., & Tillman, B. W. (1983, August). Initial test and evaluation of a computerized ability assessment technique. Paper presented at the Annual Convention of the American Psychological Association in Anaheim, CA.

Rossmeissl, P. G. Tillman, B. W. Rigg, K. E., & Best, P. R. (1984). Job assessment software system (JASS) for analysis of weapon systems personnel requirements (Research Report 1355). U.S. Army Research Institute for the Behavioral and Social Sciences, Alexandria, VA. (AD A146 448)

Wing, H., Peterson, N. G., & Hoffman, R. G. (1984, August). Expert judgments of predictor-criterion validity relationships. Paper presented at the 92nd Annual Convention of the American Psychological Association, Toronto, Canada.

APPENDIX A
DEFINITIONS OF AAS ATTRIBUTES

COGNITIVE/PERCEPTUAL

VERBAL ABILITY

THIS IS THE ABILITY TO USE AND UNDERSTAND SPOKEN AND WRITTEN LANGUAGE AND TO COMMUNICATE WITH OTHERS. IT INVOLVES "CATCHING ON" TO WHAT'S HAPPENING, COMING UP WITH AND UNDERSTANDING WORDS AND IDEAS.

MEMORY

THIS IS THE ABILITY TO MEMORIZE AND RECALL INFORMATION AND USE IT ACCORDINGLY.

REASONING ABILITY

THIS IS THE ABILITY TO THINK LOGICALLY. IT INVOLVES SEVERAL STEPS: 1) CONSIDER INFORMATION (NUMBERS, IDEAS, FACTS, OR RULES), 2) SELECT WHAT'S IMPORTANT, AND 3) PUT IT TOGETHER TO SOLVE A PROBLEM OR MAKE A DECISION. IT INVOLVES THINKING ABOUT WHY THINGS GO TOGETHER AND WHETHER THINGS "MAKE SENSE".

NUMBER FACILITY

THIS IS THE ABILITY TO ADD, SUBTRACT, MULTIPLY, AND DIVIDE QUICKLY AND CORRECTLY.

MECHANICAL COMPREHENSION

THIS IS THE ABILITY TO UNDERSTAND MECHANICAL, SHOP, AUTOMOTIVE, AND ELECTRONICS TERMS AND KNOWLEDGE, AND HOW WORKING PARTS OPERATE.

INFORMATION PROCESSING

THIS IS THE ABILITY TO ZERO IN ON NEEDED INFORMATION AND REACT WITHIN A MINIMUM AMOUNT OF TIME. INFORMATION MAY COME, FROM ONE OR MANY SOURCES. IT MAY BE NECESSARY TO SHIFT ATTENTION BACK AND FORTH BETWEEN DIFFERENT SOURCES OR TO, CONCENTRATE ON JUST ONE.

CLOSURE

THIS IS THE ABILITY TO RECOGNIZE PART-WHOLE RELATIONSHIPS. IT INCLUDES SEEING THAT SOUNDS, SHAPES, OR PIECES OF THINGS FORM A TOTAL PATTERN OR STRUCTURE (WHOLE) AND HOW THEY FIT TOGETHER. IT ALSO MEANS BEING ABLE TO LOCATE PATTERNS (PARTS) THAT ARE HIDDEN WITHIN OTHER MATERIALS.

VISUALIZATION

THIS IS THE ABILITY TO IMAGINE HOW SOMETHING WOULD LOOK. IT MAY BE SOMETHING NEVER SEEN BEFORE, OR SEEN ONLY IN A DIAGRAM OR PICTURE. OR, IT MAY BE A FAMILIAR SHAPE OR PATTERN THAT MUST BE IDENTIFIED AFTER IT IS CHANGED AROUND: BACKWARDS, UPSIDE DOWN, REVERSED, OR BELOW OTHER SHAPES.

PERCEPTUAL SPEED AND ACCURACY

THIS IS THE ABILITY TO NOTICE DETAILS ABOUT THINGS (LETTERS, NUMBERS, SOUNDS OR PATTERNS) QUICKLY AND CORRECTLY. THIS INVOLVES RAPIDLY NOTING CHANGES AND/OR THE WAY THINGS DIFFER OR ARE ALIKE.

PHYSICAL/PSYCHOMOTOR

PHYSICAL STRENGTH

THIS IS THE ABILITY TO PUSH, PULL, LIFT, AND/OR CARRY. IT MAY INCLUDE SHORT BURSTS OF EFFORT OR CONTINUOUS USE OF FORCE BY VARIOUS MUSCLE GROUPS OR THE WHOLE BODY.

STAMINA

THIS IS THE ABILITY TO MAINTAIN OR ENDURE PHYSICAL ACTIVITY OVER LONG PERIODS OF TIME WITHOUT GETTING TIRED.

MULTILIMB COORDINATION

IS IS THE ABILITY TO USE AT LEAST TWO LIMBS (ARMS, LEGS OR ARMS AND LEGS) AT THE SAME TIME.

DEXTERITY

THIS IS THE ABILITY TO MAKE SKILLFUL FINGER AND/OR HAND ACTIONS TO GRASP, PLACE OR MOVE THINGS. THESE ACTIONS MUST BE WITHIN SOME TIME LIMIT.

STEADINESS/PRECISION

THIS IS THE ABILITY TO MAKE VERY CONTROLLED BODY MOVEMENTS OR ADJUSTMENTS OF EQUIPMENT CONTROLS. IT MAY REQUIRE THINGS LIKE AIMING, SLOW, STEADY MOTIONS, "FINE TUNING" ADJUSTMENTS, AND/OR VERY FAST, VERY EXACT ACTIONS TO COUNTER CHANGES IN CONDITIONS.

NONCOGNITIVE

SOCIAL INTERACTION

THIS IS THE ATTRIBUTE THAT ENABLES PEOPLE TO BE OUTGOING AND GET ALONG WELL WITH PEOPLE INDIVIDUALLY AND IN GROUPS. IT INCLUDES WANTING TO HELP, TEACH, UNDERSTAND, AND JUST BE WITH OTHER PEOPLE.

STRESS TOLERANCE

THIS IS THE ATTRIBUTE THAT ENABLES SOMEONE TO MAINTAIN A "COOL HEAD", TO KEEP EMOTIONS UNDER CONTROL, AND TO BE PLEASANT, EASYGOING AND AGREEABLE EVEN UNDER VERY STRESSFUL CONDITIONS.

CONSCIENTIOUSNESS

THIS IS THE ATTRIBUTE THAT REFLECTS RESPECT FOR DISCIPLINE, ORDER, STRUCTURE, REGULATIONS, AND AUTHORITY. IT RESULTS IN PLANFUL, DEPENDABLE, WELL-ORGANIZED BEHAVIOR.

WORK ORIENTATION

THIS IS THE ATTRIBUTE THAT REFLECTS A BELIEF THAT HARD WORK AND PERSEVERANCE PAY OFF. IT IS CHARACTERIZED BY BELIEF THAT RESULTS ARE DUE TO ONE'S OWN EFFORTS (i.e., PERSONAL RESPONSIBILITY) RATHER THAN CHANCE EVENTS ("FATE") OR WHAT SOMEONE ELSE DOES.

SELF-ESTEEM/LEADERSHIP

THIS IS THE ATTRIBUTE THAT REFLECTS SELF-CONFIDENCE, BELIEF IN ONE'S ABILITY TO SUCCEED, AND A DESIRE TO TAKE CONTROL AND TO LEAD OTHERS. IT INCLUDES BEING FORCEFUL, PERSUASIVE, AND WILLING TO TAKE CHARGE.

ATHLETIC ABILITY/ENERGY

THIS IS THE ATTRIBUTE THAT REFLECTS TYPICALLY HIGH LEVELS OF ENERGY, ENTHUSIASM, SKILL, AND INTEREST IN TAKING PART IN PHYSICAL ACTIVITIES.

REALISTIC INTERESTS

THIS IS A PREFERENCE FOR ACTIVITIES THAT ARE PRACTICAL, CONCRETE, AND PRODUCT-ORIENTED. THESE ACTIVITIES TEND TO REQUIRE PHYSICAL, MECHANICAL AND/OR TECHNICAL SKILLS.

INVESTIGATIVE INTERESTS

THIS IS A PREFERENCE FOR SCIENTIFIC, MATHEMATICAL, OR INTELLECTUAL ACTIVITIES. THESE INVOLVE THINKING AND ORGANIZING, OBSERVING, ANALYZING, EVALUATING, AND/OR TESTING PRODUCTS OR IDEAS.

APPENDIX B

SAMPLE PAGE: FIRST ANCHOR RATING INSTRUMENT

VERBAL ABILITY

THIS IS THE ABILITY TO USE AND UNDERSTAND SPOKEN AND WRITTEN LANGUAGE AND TO COMMUNICATE WITH OTHERS. IT INVOLVES "CATCHING ON" TO WHAT'S HAPPENING, COMING UP WITH AND UNDERSTANDING WORDS AND IDEAS.

		1	2	3	4	5	6	7
1. UNDERSTAND SIMPLE SAFETY SIGNS.	1.							
2. WRITE UP AN ACCIDENT REPORT GIVING ALL THE IMPORTANT INFORMATION.	2.							
3. WRITE A TECHNICAL MANUAL ON HOW TO PERFORM YOUR JOB.	3.							
4. REPORT INFORMATION ABOUT ENEMY TROOPS USING 'SALUTE'.	4.							
5. PREPARE A WRITTEN SUMMARY OF THE TRAINING YOU GOT IN A SPECIAL COURSE.	5.							
6. EXPLAIN PROCEDURES TO FOLLOW TO GIVE FIRST AID TO BURN VICTIMS.	6.							
7. WRITE UP DAILY REPORTS OF YOUR UNIT'S OPERATIONS.	7.							
8. EXPLAIN CONCEPTS OF AN OPERATION TO OTHER SOLDIERS IN SEVERAL DIFFERENT WAYS.	8.							
9. LEAD AND UNDERSTAND AN OPERATIONS ORDER.	9.							
10. PREPARE AN EQUIPMENT REQUISITION ORDER.	10.							

APPENDIX C

SAMPLE PAGE: SECOND ANCHOR RATING INSTRUMENT

STRESS TOLERANCE

THIS THE THE ATTRIBUTE THAT ENABLES SOMEONE TO MAINTAIN A "COOL HEAD", TO KEEP EMOTIONS UNDER CONTROL, AND TO BE PLEASANT, EASYGOING AND AGREEABLE EVEN UNDER VERY STRESSFUL CONDITIONS.

HOW MUCH OF THIS ATTRIBUTE IS NEEDED FOR SOMEONE TO DO THE FOLLOWING THINGS?

DISREGARD ANY TRAINING OR OTHER ABILITIES OR ATTRIBUTES THAT MAY BE INVOLVED.

_____ 1. STAY CALM WHEN SOMEONE CRASHES INTO THEIR VEHICLE.

_____ 2. EFFECTIVELY TAKE MEASURES WHEN THEY HEAR "INCOMING."

_____ 3. ACCEPT THE FACT THAT EVERYONE PULLS EXTRA DUTY WHEN OTHER PEOPLE SHOW UP LATE.

_____ 4. DIRECT CIVILIAN VISITORS TO SAFETY WHEN A FIRE BREAKS OUR IN THE BUILDING THEY'RE IN.

_____ 5. APPLY APPROPRIATE FIRST AID WHEN THEIR BEST FRIEND IS SERIOUSLY WOUNDED BY ENEMY FIRE.

_____ 6. MAINTAIN A CHEERFUL ATTITUDE WHEN THEIR LEAVE IS CANCELLED FOR GOOD REASON.

_____ 7. ORGANIZE RESCUE OPERATIONS IN A SEVERE EMERGENCY.

_____ 8. COOPERATE WHEN TOLD TO DO SOME DISTASTEFUL TASK THAT NORMALLY ISN'T PART OF THEIR JOB.

_____ 9. DON'T GRIPE WHEN THEY'RE SENT TO THE FIELD SEVEN WEEKS OUT OF EIGHT.

_____ 10. ADJUST TO ASSIGNMENT TO A NEW POST.

/ 1 / 2 / 3 / 4 / 5 / 6 / 7

very small amount	moderate amount	very great amount

WHICH ITEM NEEDS THE GREATEST AMOUNT OF THIS ATTRIBUTE? _____
WHICH ITEM NEEDS THE LEAST AMOUNT OF THIS ATTRIBUTE? _____

GO BACK AND REVIEW YOUR RATINGS. DO NOT GO ON TO THE NEXT PAGE YET.

APPENDIX D

SAMPLE PAGE: ATTRIBUTE ASSESSMENT FROM EXPERIMENT 2

VISUALIZATION

THIS IS THE ABILITY TO IMAGINE HOW SOMETHING WOULD LOOK. IT MAY BE SOMETHING NEVER SEEN BEFORE, OR SEEN ONLY IN A DIAGRAM OR PICTURE. OR, IT MAY BE A FAMILIAR SHAPE OR PATTERN THAT MUST BE IDENTIFIED AFTER IT IS CHANGED AROUND: BACKWARDS, UPSIDE DOWN, REVERSED, ABOVE OR BELOW OTHER SHAPES.

7

6

-- ASSEMBLE A RADIO FROM A DIAGRAM IN A MANUAL

5

4 -- MENTALLY PICTURE WHERE TO PLACE TACTICAL EQUIPMENT AND MATERIALS IN THE FIELD

3

-- READ A WRISTWATCH UPSIDE DOWN

2

1

Ø -- REQUIRES NONE AT ALL