

AD A 076826

LEVEL II

Research Memorandum 76-28

(25)

**RELIABILITY OF ASSOCIATE RATINGS OF
PERFORMANCE POTENTIAL BY
ARMY AVIATORS**

Robert F. Eastman and Robert L. McMullen



FORT RUCKER FIELD UNIT

DDC FILE COPY



This document has been approved
for public release and sale; its
distribution is unlimited.

U. S. Army

Research Institute for the Behavioral and Social Sciences

November 1976

79 11 13 305

DISPOSITION FORM

For use of this form, see AR 340-15, the proponent agency is TAGCEN.

REFERENCE OR OFFICE SYMBOL	SUBJECT
PERI-TP	Clearance and Transmittal of Reports to DTIC

TO DDC-DAA-1
ATTN: Mr. Schrecengost

FROM ARI Rsch Pub Group

DATE 8 Nov 79

CMT 1

Ms Price/48913

1. The reports listed on Inclosure 1 are approved for public release with unlimited distribution (50 numbered ARI Research Memorandums, 74-1 thru 76-30).
2. These are among the previously unrecorded ARI reports which you identified to us 22 June 1979 as not in your retrieval system. The accompanying box contains at least one copy of each report for your retention and reproduction.

1 incl
List of reports, 1974-76

Helene S. Price

HELEN S. PRICE
Research Publications Group
Army Research Institute

NTIS	1
DTIC	1
DDC	1
DDI	1
DDP	1
DDT	1
DDU	1
DDV	1
DDW	1
DDX	1
DDY	1
DDZ	1
Dist.	1
A	1

Army Project Number

(16) 20763743A772

Aviator Selection

(9) *memois*
Research Memorandum 76-28

(6) RELIABILITY OF ASSOCIATE RATINGS OF PERFORMANCE
POTENTIAL BY ARMY AVIATORS

(10) Robert F. Eastman Robert L. McMullen

Submitted by:
Charles A. Gainer, Chief
Fort Rucker Field Unit

(12) 12/

(14)

ARI-BM-76-28

(11)

Nov 1976

Approved by:

Joseph Zeidner, Director
Organizations and Systems Research
Laboratory

J. E. Uhlaner, Technical Director
U.S. Army Research Institute for the
Behavioral and Social Sciences

Research Memorandums are informal reports on technical research problems. Limited distribution is made, primarily to personnel engaged in research for the Army Research Institute.

408 010

xlh

RELIABILITY OF ASSOCIATE RATINGS OF PERFORMANCE POTENTIAL BY ARMY AVIATORS

BACKGROUND

In response to a TRADOC request, the Fort Rucker Field Unit of the Army Research Institute for the Behavioral and Social Sciences has undertaken a program to determine the attributes which make some attack pilots outstanding performers in combat. The program which was proposed consists of the following three inter-related tasks: (1) development of an attack pilot profile from survey data of proven performers; (2) development of a rating form to evaluate potential attack pilots; and (3) evaluation and assessment of AH-1 trainees¹ using the findings of tasks 1 and 2.

Currently no systematic selection of candidates for AH-1 transition exists. Many trainees do not want to be attack pilots but are assigned to fill requirements because they are due for reassignment. If unit commanders and personnel had more information, a better fit of aviators to transition training might be possible. This research is part of task 2 and was done to determine the reliability of unit-level ratings of potential AH-1 candidates.

OBJECTIVES

The primary objectives of this study were to evaluate the inter-rater reliability of a rating form designed to select AH-1 transition candidates and to study rater acceptance of rating forms.

It was hypothesized that COBRA pilots in cavalry and assault units would demonstrate high inter-rater reliability when rating non-AH-1 qualified aviators in their units on potential for AH-1 transition and gunship pilot duties. A variable of interest was the effect of length of rater-tee acquaintance on inter-rater reliability (Freeberg, 1969). —→ p. 5

Structured interviews of more than 50 attack pilots revealed that about 70% felt that they could predict other pilots who were potential "ACEs" (i.e., attack pilots who perform exceptionally well in combat) without observing them under combat conditions. This preliminary research suggested that a majority of attack pilots would favor use of rating forms to evaluate candidates for AH-1 transition in contrast to a general lack of acceptance of ratings to determine leadership potential in Army schools (Medland, Yates and Downey, 1974; Downey, 1975).

¹ AH, attack helicopter = COBRA; UH, utility helicopter.

METHOD

SAMPLE

Ratees. The ratees were UH-1 and UH-1/OH-58 qualified aviators in six troops of the 6th ACCB, one troop of the 2d Armor Division, and one troop of the 1st Cavalry. For some units all eligible ratees were used. For those troops which had more than 13 eligible ratees, a random procedure was used to eliminate individuals and reduce the number of ratees to 12. The modal number of ratees was 12 (see Table 1). Because of anonymity requirements imposed by unit commanders, no analysis of the backgrounds of ratees was possible.

Raters. The raters consisted of all the available AH-1 qualified aviators in the eight aviation units. There was considerable variation in the number of raters available within each troop (see Table 1). Because of field duty assignments many AH-1 aviators were not available for participation in this study. However, no systematic basis for non-availability which would affect this study was apparent.

PROCEDURE

Several troop commanders were unwilling to have their personnel participate in this study unless careful procedures were used to insure the anonymity of raters and ratees. This requirement was satisfied as follows: The raters in each troop were scheduled to arrive at the classroom as a group. Each rater within a group was given a packet with a randomly preassigned two-digit number as he entered the classroom. When the entire troop arrived, the list of names of ratees for the troop was placed on the blackboard and a set of three-digit numbers were randomly assigned to each ratee. The raters were told that the set of numbers on the board corresponded to three-digit numbers appearing on the evaluation forms in their packets. It was clearly stated that no previous connection between names and numbers had existed. Raters were then instructed to consider the entire list of individuals before following the written instructions on the rating forms. Immediately after each troop completed their ratings the names and numbers on the blackboard were erased. No records of ratees' numbers were kept.

RATING SCALE

The rating scale was modeled after a general format used at several Army schools² and is designed to encourage discrimination among ratees

² For example, The RANGER Evaluation Report; FB (RD) Form 2.

Table 1
SIGNIFICANCE OF OBSERVED INTER-RATER AGREEMENT AMONG RATERS--COEFFICIENTS
OF CONCORDANCE (W) AND AVERAGE RHO (\bar{r}_s)

Unit No.	Troop Designation	Number of Raters	Number of Ratees	W	\bar{r}_s	χ^2	P <
1	B 7/17	11	12	.381	.32	46.049	.001
2	C 4/9	6	12	.541	.45	35.706	.001
3	C 7/17	10	11	.502	.45	50.200	.001
4	A 7/17	10	12	.559	.51	61.490	.001
5	D 1/9	6	12	.378	.25	24.948	.01
6	D 2/1	6	13	.342	.21	24.624	.02
7	HHT 4/9	3	13	.724	.58	26.064	.02
8	A 4/9	7	12	.349	.24	26.870	.01
Total		59	97				

Note. $\bar{r}_s = \frac{mW-1}{m-1}$ where m = number of raters (Hayes, 1973).

on a set of desirable characteristics for gunship/attack pilots. The attack pilot characteristics were identified during structured interviews of attack pilots with combat experience. On the evaluation form the rater is instructed to consider the attack pilot characteristics and assign the ratee a numerical rank (i.e., 1-25 in a typical group of 25 pilots) representing potential for success as an attack pilot. The rater is also provided space for a 2-3 sentence word picture to justify the numerical rating assigned. Additional information is recorded on where the rating was conducted and the type and duration of the relationship between rater and ratee. Detailed instructions also spell out the restrictions to be followed when rating a group of AH-1 candidates (see Appendix A).

RATING FORM EVALUATION

The rating form evaluation was designed to ask AH-1 qualified raters the following: (1) characteristics to add to or delete from the list on the rating form, (2) if ratings should be used and who should do the rating, and (3) open-ended questions about why they do or don't favor ratings and potential operational problems of using ratings.

RESULTS AND DISCUSSION

Coefficients of concordance and average rho's were computed for each unit to determine the degree of agreement among raters. The results in Table 1 indicate that inter-rater agreement was very high ($p < .01$) for 6 of the 8 units studied. Although statistical significance was not quite achieved in units 6 and 7 ($p < .02$), the number of raters involved was small. The overall results indicate that raters were responding to essentially the same characteristics when rating the candidates in their units. This does not mean that the observed orderings were necessarily correct, or based on the "true" criteria for a successful gunship pilot; however, the location of the attack pilot characteristics on the rating form itself was designed to make the rating criteria explicit and reduce the influence of irrelevant variables.

The results of the Rating Form Evaluation are presented in Appendix B. The percentage of the sample (if over 5%) who responded to an alternative or who made an open-ended comment is shown on a copy of the rating form itself. The most important result was that 58% of the respondents felt ratings should be used in selecting candidates for COBRA transition training. However, many raters indicated reservations about the validity of ratings because of the influence of personal prejudices and friendships (20%) and the adequacy of ratings to assess candidates effectively (15%). To the question "who do you think should rate AH-1 candidates," only 20% of the sample responded "commanding and supervisory personnel," while 58% indicated either AH-1 qualified pilots (29%) or other pilots in the unit (29%).

Another variable which is important in peer rating research is how well the rater knows the individual being rated. The measure of association recorded in this study was the length of time the rater knew the rated officer. The relationship between median duration of rater-ratee acquaintance within each unit and the statistical reliability of inter-rater agreement is shown in Table 2. The data in Table 2 suggest that inter-rater reliability, in these units, may be positively related to the duration of association between raters and ratees.

Table 2
RELATIONSHIP BETWEEN INTER-RATER RELIABILITY AND
DEGREE OF RATER-RATEE ACQUAINTANCE WITHIN UNITS

Unit No.	Median Number of Months of Rater- Ratee Acquaintance	Significance Level of Inter-Rater Agreement
1	18	$p < .001$
2	12	$p < .001$
3	12	$p < .001$
4	15	$p < .001$
5	6	$p < .01$
6	9	$p < .02$
7	10	$p < .02$
8	6	$p < .01$

CONCLUSIONS

AH-1 qualified aviators demonstrate a high degree of inter-rater reliability when rating AH-1 candidates in their units. Statistical reliability was higher within those units where raters and ratees had been acquainted longer.

The majority (58%) of AH-1 qualified raters sampled feel that ratings by fellow aviators should be used in selecting candidates for COBRA transition.

Many raters (20%) had reservations about the use of ratings because of the influence of friendship and other personal biases. Another group felt ratings could not (8%) or might not (7%) be effective in identifying potentially successful AH-1 candidates.

REFERENCES

Downey, R. G. Associate evaluations: Improving field acceptance. ARI Research Memorandum 75-5. July 1975.

Freeberg, N. E. Relevance of rater-ratee acquaintance in the validity and reliability of ratings. Journal of Applied Psychology, 1969, 53, 512-524.

Hayes, W. L. Statistics for the social sciences (2d ed.). New York: Holt, Rinehart and Winston, 1973.

Medland, F. F., Yates, L. G. and Downey, R. G. Associate ratings of senior officer potential. ARI Research Problem Review 74-2. June 1974.

APPENDIXES

Appendix	Page
A. ATTACK PILOT CANDIDATE EVALUATION FORM	11
B. RATING FORM EVALUATION	13

APPENDIX A

ATTACK PILOT CANDIDATE EVALUATION FORM

Instructions:

1. Evaluate this man in your unit/class in terms of your estimate of his potential ability to become a successful gunship/attack pilot. Determine where you think he would rank in a typical group of 25 pilots (number 1 the highest ranking, 25 the lowest ranking). Consider the ATTACK PILOT CHARACTERISTICS below prior to rating each man. Consider the entire group you are asked to evaluate and the following restrictions before beginning. (a) No more than two individuals may be placed in 1-5 column. (b) no two individuals will be assigned the same rating number. Do not rate yourself.
2. Under REMARKS, write a 2-3 sentence word picture to justify the numerical rating you assigned. State briefly the characteristics (desirable or undesirable) of this man that impressed you most.
3. Your ratings will remain anonymous. The packet you picked up has an ID number only to insure that you followed the restrictions when rating.

EVALUATED INDIVIDUAL'S NAME (Last, first)				DATE DAY MONTH YEAR / /				
ATTACK PILOT CHARACTERISTICS								
TACTICAL KNOWLEDGE	AGGRESSIVENESS			CONFIDENCE				
PHYSICAL STAMINA	SELF-DISCIPLINE			TEAMWORK				
TIMELINESS OF ACTION	DRIVE			INITIATIVE				
COMBAT INSTINCT	EFFECTIVE MAP USE			DEPENDABILITY				
EVALUATED PERSON'S PRESENT LOCATION. (Circle one)	FW	UNIT	TRANSITION TRAINING	STANDING WITHIN A 25-MAN GROUP (Circle one)				
RELATIONSHIP TO EVALUATED PERSON (Circle one)	CO	IP	OTHER	1	6	11	16	21
REMARKS:				2	7	12	17	22
				3	8	13	18	23
				4	9	14	19	24
				5	10	15	20	25
HOW LONG HAVE YOU KNOWN THE INDIVIDUAL? _____								
RATER ID # _____								

APPENDIX B

RATING FORM EVALUATION

1. What characteristics would you add to the list on the rating form?

% Responding

Desire	17%
Common Sense	17%
Mental Stamina	7%
Maturity	5%

2. What characteristics would you delete from the list on the rating form?

% Responding

Combat Instinct	14%
Physical Stamina	8%

3. Who do you think should rate AH-1 candidates?

AH-1 qualified pilots only	All other pilots in their unit	Commanding and supervisory personnel only	Instructor pilots only	Senior aviators only
----------------------------	--------------------------------	---	------------------------	----------------------

% Responding

(29%)	(29%)	(20%)	(14%)	(8%)
-------	-------	-------	-------	------

4. What problems, if any, can you foresee in using a rating form like the Attack Pilot Candidate Evaluation form at the unit level? Use the reverse side if more space is needed.

% Responding

Personal prejudices and friendships will invalidate it	20%
Form can't judge qualifications	15%

5. Do you feel that ratings by fellow aviators should be used in selecting candidates for COBRA transition?

YES	NO
% Responding (58%)	(42%)

Briefly indicate in the space below why you feel ratings should/should not be used. Use the reverse side if more space is needed.

<u>Should</u>	<u>% Responding</u>	<u>Should Not</u>	<u>% Responding</u>
Fellow aviators more aware of abilities	12%	Ratings are biased	8%
		Ratings are no good	8%
		Ratings may not be any good	7%