Professional Paper 34-69
November 1969

Combat Aviator Criterion Development

by

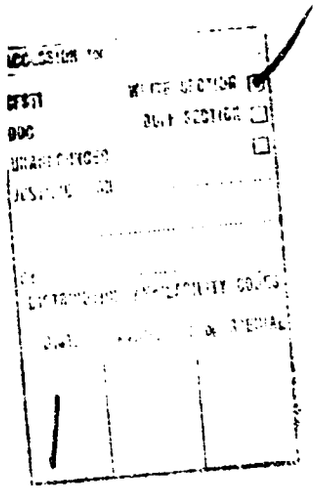
Wiley R. Boyles, Peter R. Prunkl,
and James L. Wahlberg

Symposium Presentation at the
American Psychological Association
Convention
Washington, D.C. September 1969

HumRRO

HUMAN RESOURCES RESEARCH ORGANIZATION

15



The Human Resources Research Organization (HumRRO) is a nonprofit corporation established in 1969 to conduct research in the field of training and education. It is a continuation of The George Washington University Human Resources Research Office. HumRRO's general purpose is to improve human performance, particularly in organizational settings, through behavioral and social science research, development, and consultation. HumRRO's mission in work performed under contract with the Department of the Army is to conduct research in the fields of training, motivation, and leadership.

Published
November 1969
by
HUMAN RESOURCES RESEARCH ORGANIZATION
300 North Washington Street
Alexandria, Virginia 22314

Prefatory Note

This paper was presented by Dr. Boyles as part of a tri-service symposium entitled, *Development of a Combat Aviator Criterion.* Other contributors to the symposium represented the Naval Aerospace Medical Institute, Pensacola, Florida, and the U.S. Air Force Human Resources Laboratory.

This paper describes research performed by Division No. 6 (Aviation), of the Human Resources Research Organization, at Fort Rucker, Alabama, under Work Unit PREDICT, Correlational Analysis of Aviator Performance. The paper deals with one aspect of the Work Unit, combat aviator criterion development.

COMBAT AVIATOR CRITERION DEVELOPMENT

Wiley R. Boyles, Peter R. Prunkl,
and James L. Wahlberg

Many different methods for evaluating the effectiveness of human work behavior are found in the professional literature. Two of the major methods are objective measures of performance—in terms of what Dunnette (1) calls "organizational consequences"—and ratings by persons who have had opportunity to observe worker performance. Of the rating methods, graphic rating scales are most common; others include critical-incident reports, forced-choice techniques, free-written ratings, nominations (by peers, subordinates, and superiors), ranking methods, checklists, and work-sample tests (2). Appropriateness of technique is a function of the complexity of the work, the types of raters available, the structure of the organization where the evaluations are to be used, the purpose of the evaluations, and to use Dunnette's (1) terms again, the "situational and social circumstances" of the job environment.

In relating "appropriateness" to the Army aviation combat situation, we are faced with the necessity of evaluating performance on highly complex jobs in work situations in which many types of raters are available. Traditionally, the Army has used ratings by superiors as a means of evaluating job performance in operational situations. We are interested in obtaining evaluations of job performance for purposes of research, intending to use them for validation of selection and training techniques. Also, the validation must be in terms of multiple criteria, for our aviators must be effective officers as well as effective pilots. The situational and social circumstances of the jobs range widely. Objective performance measures, in terms of organizational consequences, are difficult to collect.

The Army's Behavioral Science Research Laboratory has launched a systematic effort to establish criteria for *officer* combat performance. However, no large-scale Army effort to evaluate potential criteria of *aviator* performance in combat now exists. HumRRO has taken some preliminary steps in this direction, and further steps are planned. We welcome technical suggestions. We have surveyed the literature, and have some data that show promise, but at this time we have many more questions than answers.

At the HumRRO Aviation Division, we are constructing a data bank for Army aviators. We are working toward the establishment of a multivariate system for prediction of performance in both training and operational situations. Criteria of training success are relatively easy to establish; criteria of individual adequacy in Army aviation combat obviously are not.

Our research has included, over the past two years, the administration of paper-and-pencil attitude measures developed in earlier HumRRO research on motivation under combat stress. Because these inventories seemed to have promise for aviation training and secondary selection situations, we have administered them to several thousand potential aviators. Our results show that the inventories differentiate between aviator candidates and other members of the basic combat trainee population, and between aviation trainees who successfully complete flight training and those who fail because of flying deficiencies.

Concurrently with administration of these inventories, we began exploratory evaluations of problems in combat criterion development. Rickus and Berkshire (3) discussed in 1968 the problems encountered during past efforts to evaluate the performance of the combat aviator. Some of these problems are results of the complexity of combat; others are functions of the complexity of aviation.

In our first step toward deriving criteria for combat validation of the data bank material, the critical incident technique to collect samples of ineffective behavior was used. This research was reported in 1968 (4). Among the results was a finding similar to that of the Navy; refusal of a pilot to fly with another aviator is a frequent identifying incident connected with perceived ineffectiveness. The results also indicated that a manageably small number of behavior categories accounted for a large portion of the critical incidents.

Building on these results, we have recently begun to collect data with an aviator combat performance rating instrument that includes peer nominations, a number of five-point graphic rating scales, and some dichotomous items. Some of the scales and items were based on our critical incident data, and some were based on subsequent interviews with veteran combat aviators. Others were taken from instruments developed by Roff (5) and by Rickus and Berkshire (3). An Appendix to this paper contains the nomination format, rating scales, and dichotomous items chosen from a pretest. In the Rickus and Berkshire study, Navy aviators agreed that the flight surgeon was in the best position to rate their combat performance. However, in our work, interviews, questionnaires, and consensus of aviators indicated that fellow aviators are in the best position to evaluate Army aviator combat performance.

We are assessing the psychometric properties of our items in terms of their ability to distinguish between aviators nominated as "most proficient" or "least proficient" by peers, subordinates, and superiors. Since there are two distinct categories of Army aviators—warrant officers and commissioned officers—we are exploring differences in ratings as a function of category interrelationships. Additional possible moderator variables included are aviation unit size, unit mission, terrain types over which an aviation unit operates (flying behavior requirements range widely as a function of terrain), and frequency of observation.

In general, our scales deal with quality and speed of combat decisions, flexibility, combat flying content knowledge, risk-taking behavior, confidence, leadership, endurance, and emotional stability.

One objective performance scale item, relating to selection by superior for tactical emergency missions, was included. The dichotomous items call for some responses in terms of objective performance measures and others in terms of subjective evaluations. As would be expected from their content, almost all produce significant χ^2s between least proficient and most proficient nominees. The two exceptions are "turned in wings" and "killed in action" which are events too rare to produce significance. The proportion of *wounded* in action is significantly higher for those aviators nominated as "most proficient" $\{\chi^2 = 17.60, p<.001\}$.

As in the 1968 Rickus and Berkshire study (3), the most frequently cited identifying behavior for low proficiency aviators was "people refuse to fly with him." This was cited for 61% of the aviators nominated as "least proficient," while reported for less than 1% of those named "most proficient." On the other hand, the item which appeared with second highest frequency in the Rickus and Berkshire study, "turned in his wings," was encountered in only two of 201 cases in our recently collected data. The item on the HumRRO scale that appeared with second highest frequency among least proficient aviators was item 26, "Did he react to loud noises in the aircraft as if they were gunshots?" While 50% of the least proficient aviators were reported to have reacted in this manner, 20% of the most proficient aviators were also reported to have reacted this way. Thus, while this item was significant at the .001 level, there were other items that were better discriminators between most and least proficient aviators. For example, 47% of the least proficient aviators were said to panic in "hot" situations, whereas less than 1% of the most proficient aviators did so.

Let me note that we asked this sample for nominations of aviators "least proficient while flying under enemy fire," rather than for "ineffective" or "unsatisfactory" aviators. We are dependent for definition of a satisfactory-unsatisfactory dichotomy on responses to Item 22 of our scale. Item 22 results showed 55 of 201 aviators nominated as "least proficient" were reported as having been removed from the job because of poor performance.

Of our situational context variables, almost all proved to be significant moderators of one or more items. This fact warrants early cross-validation of results.

Asymmetric moderation was obvious in the data. For example, we found no significant difference between the frequencies with which warrant officers and commissioned officers were named "least proficient" as a function of the rater's category. That is, commissioned officers seemed no more prone to name warrant officers "least proficient" than did the warrant officers themselves. On the "most proficient" item, however, the situation was quite different. Warrant officers more frequently named warrant officers "most proficient;" commissioned officers named commissioned officers $\{\chi^2 = 12.00, p<.001\}$. This sort of finding emphasizes the necessity for studying all possible moderator combinations in the data.

Apart from the rating scale data, two other sources offered information on our aviators: a Background Activities Inventory (BAI), the previously mentioned attitude inventory designed from a model of reaction to combat stress, and a Vietnam Returnee Questionnaire administered to all aviators returning to Fort Rucker from Vietnam.

From the BAI, we found that age upon entering the flight program was significantly higher for those aviators named "most proficient"

$$\{t = 2.40,\ p<.05;\ r_{ph} = .40,\ p<.01\}.$$

The mean age was 21.2 years for this group as compared to 19.9 years for the "least proficient" group.

These results were not a function of differences in amount of combat experience. The 47 pilots in this sample had one tour in Vietnam as aviators and had approximately the same amount of experience.

The BAI also furnished information regarding past participation in dangerous sports and activities. We hypothesized that our pilots rated "most proficient" would have participated more often in these activities than those rated "least proficient." Our data did not support this hypothesis: The most proficient aviators had significantly lower exposure to these sports and activities than the least proficient pilots

$$\{t = 2.60,\ p<.05;\ r_{ph} = -.35,\ p<.05\}.$$

From our second additional data source, the Vietnam Returnee Questionnaire, various self-report items completed by 21 of our ratees were selected. No items showed a significant difference between the two groups. There was no significant difference in self-reports of days flown in an average month, number of accidents, number of forced landings, and number of times grounded.

I have given a rather cursory look at the details of our results, as this presentation is intended to be more provocative than documentary. I think we are getting some thought-provoking data. In answer to "Where do we go from here?" I will list some of the alternatives under consideration. One is a refinement of our scales using the method developed by Smith and Kendall (6), which seems to have great potential for optimizing the combination of the components of psychometric soundness on the one hand, and user acceptability and accuracy on the other. That method proceeds from critical incident collections, categorized by major qualities necessary for job success, through independent matching of qualities and incidents, to independent judgmental locations of the instruments on a scale, with all steps performed by content experts and with vigorous requirements for acceptability of both the behavior descriptions making up the scales and qualities measured by the scales.

Another possibility is exploration of the utility of the multitrait, multirater method, either in combination with or as an alternative to the Smith and Kendall (6) approach. The utility of this approach, as described by Lawler (7), seems to hold up well across a variety of complex jobs.

Finally, I would like to emphasize the continuing importance of systematic pan-service coordination of behavioral studies. Both significant overlap and important differences exist among critical job requirements for Navy, Marine, Air Force, Coast Guard, and Army aviators. Comparison of our preliminary data in this area with those of Rickus and Berkshire (3) shows both strong convergence and strong divergence. Knowledge of the commonalities and disparities in our aviator job requirements could have a synergistic effect.

LITERATURE CITED

1. Dunnette, Marvin D. *Personnel Selection and Placement*, Wadsworth Publishing Company, Inc., Belmont, Calif., 1966.

2. Barrett, Richard S. *Performance Rating*, Science Research Associates, Inc., Chicago, 1966.

3. Rickus, George M., Jr., and Berkshire, James R. *Development of an Aviation Combat Criterion: Preliminary Report*, NAMI-1047, Naval Aerospace Medical Institute, Naval Aerospace Medical Center, Pensacola, Fla., August 1968.

   Prunkl, Peter R., and Boyles, Wiley R. *A Preliminary Application of the Critical Incident Technique to Combat Performance of Army Aviators*, presentation at the Alabama Psychological Association annual meeting, Birmingham, Ala., May 1968, issued as HumRRO Professional Paper 24-68, June 1968.

5. Roff, Merrill F. *A Study of Combat Leadership in the Air Force by Use of a Rating Scale*, Project Number 21-02-013, USAF School of Aviation Medicine, Randolph Field, Tex., February 1950.

6. Smith, Patricia, and Kendall, L.M. "Retranslation of Expectations: An Approach to the Construction of Unambiguous Anchors for Rating Scales," *J. Appl. Psychol.*, vol. 47, 1963, pp. 149-155.

7. Lawler, Edward E., III. "The Multitrait-Multirater Approach to Measuring Managerial Job Performance," *J. Appl. Psychol.*, vol. 51, no. 5, Part I, October 1967.

Appendix

COMBAT RATING SCALE

A. UNIT INFORMATION

1. Your name and grade: _____
2. Name the unit with which you served the longest during your last RVN tour. _____
3. Name the date that you reported to this unit. _____
4. Name the date that you left this unit. _____
5. How many aviators were assigned to this unit during your tour? _____
6. Primarily, what type of mission(s) did you fly in this unit? _____

B. AVIATOR INFORMATION

1. Think of the aviators in your unit who you have observed frequently under enemy fire. Name three of these men who you consider most proficient in their performance while under fire. (Use 1 as most proficient, 2 as second most proficient, etc. Feel free to use any aviators in your unit—superiors, subordinates, or any others—living or dead.)

   1. _____
   2. _____
   3. _____

2. Name the three aviators you consider the least proficient while flying under enemy fire. (Use 1 as least proficient, 2 as second least proficient, etc. Again you may use any aviators in your unit—superiors, subordinates, or any others—living or dead.)

   1. _____
   2. _____
   3. _____

C. CHARACTERISTIC INFORMATION

You have just rated six people in terms of their flying proficiency in combat situations. Now rate each of the six aviators you have named in terms of the following characteristics.

Remember that you are rating each person at a time on each characteristic. A good combat aviator may have some bad points and a bad aviator may have several good points while flying under enemy fire.

RECORD YOUR ANSWERS ON THE ANSWER SHEETS. Read each characteristic carefully. A is the extreme and corresponds to the behavior directly below A. E is the opposite extreme and its description is also below the letter. Consider C to be the average behavior. In some cases the behavior characteristic for C is written below it.

RATING SCALE

	A	B	C	D	E
1.	Decisions always worked when under enemy fire				Decisions never worked when under enemy fire
2.	Indecisive in his combat decisions				Reached a decision rapidly when under fire
3.	Flew mechanically in "hot" situations				Adjusted to the changing tactical situations
4.	Had learned the tricks of combat flying				Little knowledge of combat flying
5.	Excellent flying judgment while under fire				Extremely poor judgment while under fire
6.	Remembered his training in "hot" situations				Forgot important procedures while under enemy fire
7.	Took no chances while in combat				Took many chances while under fire

	A	B	C	D	E
8.	Calm and relaxed under fire				Extremely nervous and tense under fire
9.	Not confident about his ability when under fire		Confident, he felt he could handle any situation		Overconfident about his ability when under fire
10.	Was unwilling to fly tactical emergency missions				Was willing to fly the tactical emergency missions
11.	Was usually picked by superior for tactical emergency missions				Was never picked by superior for tactical emergency missions
12.	Practiced the principles of good leadership				Never practiced the principles of good leadership
13.	Sought out leadership situations				Avoided leadership situations
14.	Did not do his job when under fire		Did his job and nothing more when under fire		Did more than what was expected of him when under fire
15.	When under enemy fire flew at unnecessarily high altitude		Flew at appropriate altitude when under fire		Flew unnecessarily close to the target when under fire
16.	Always kept a clear head under fire				"Lost his head" in combat situasituations

	A	B	C	D	E
17.	Fatigued easily when under fire				Had great endurance under fire

	A	B	C	D	E
18.	Evidenced feelings of fear before difficult missions				Never showed his feelings of fear before difficult missions

D. SUMMARY INFORMATION

19. Did he ever turn his wings in?
20. Did he ever want to turn his wings in?
21. Did people ever refuse to fly with him?
22. Was he given nonflying duties because of poor performance?
23. Did he turn down aircraft that other aviators felt were flyable?
24. Was he ever charged with one or more accidents?
25. Did he ever appear before a flight evaluation board?
26. Did he react to loud noises in the aircraft as if they were gunshots?
27. Did he claim to be sick in order to keep from flying?
28. Did he ever abort a mission unnecessarily?
29. Did he panic in "hot" situations?
30. Did he frequently downgrade Army aviation?
31. Was he ever wounded in action?
32. Was he killed in action?

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)* Human Resources Research Organization (HumRRO) 300 North Washington Street Alexandria, Virginia 22314	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

COMBAT AVIATOR CRITERION DEVELOPMENT

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

Professional Paper

5. AUTHOR(S) *(First name, middle initial, last name)*

Wiley R. Boyles, Peter R. Prunkl,
and James L. Wahlberg

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
November 1969	13	7

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
DAHC 19-70-C-0012 b. PROJECT NO. 2Q062107A712 c. d.	Professional Paper 34-69 9b. OTHER REPORT NO.(S) *(Any other numbers that may be assigned this report)*

10. DISTRIBUTION STATEMENT

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

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
Presentation at American Psychological Association convention, Washington, D.C. September 1969. Work Unit PREDICT	Office, Chief of Research and Development Department of the Army Washington, D.C. 20310

13. ABSTRACT

Factors that must be considered in the development of criteria for proficient performance of a complex job are discussed in the context of the Army aviation combat situation. Ratings of aviators by peers, subordinates, and superiors on pertinent job behaviors have been collected following identification of the pertinent behaviors by the critical incident method. Moderator variable effects on the ratings are described, as are some alternative strategies for future research.

DD FORM 1 NOV 68 1473

Unclassified

Security Classification

Unclassified

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Army Aviators						
Attitude Inventories						
Combat Aviation						
Criterion Development						
Performance Prediction						
Rating Scales						

Unclassified

Security Classification