AL/HR-TP-1995-0008

DIMENSIONS OF AIR FORCE PILOT COMBAT PERFORMANCE

Michael W. Murray, 1Lt , USAF
Frederick W. Siem
Anne P. Duke
Joseph L. Weeks

HUMAN RESOURCES DIRECTORATE
MANPOWER AND PERSONNEL RESEARCH DIVISION
7909 Lindbergh Drive
Brooks AFB, Texas 78235-5352

19950811 085

August 1995

Final Technical Paper for Period November 1992 - April 1994

ARMSTRONG LABORATORY

AIR FORCE MATERIEL COMMAND
BROOKS AIR FORCE BASE, TEXAS

NOTICES

Publication of this paper does not constitute approval or disapproval of the ideas or findings. It is published in the interest of scientific and technical information (STINFO) exchange.

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility or any obligation whatsoever. The fact that the Government may have formulated or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication, or otherwise in any manner construed, as licensing the holder, or any other person or corporation; or as conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The Office of Public Affairs has reviewed this paper, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This paper has been reviewed and is approved for publication.

MICHAEL W. MURRAY, 1LT, USAF
Behavioral Scientist
Aircrew Selection Research Branch

PATRICK C. KYLLONEN
Technical Director
Manpower & Personnel Research Div

GARY D. ZANK, Colonel, USAF
Chief
Manpower and Personnel Research Division

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	August 1995	Interim - November 1992 - April 1994

4. TITLE AND SUBTITLE

Dimensions of Air Force Pilot Combat Performance

5. FUNDING NUMBERS

PE - 62205F
PR - 7719
TA - 25
WU - 01

6. AUTHOR(S)

Michael W. Murray
Frederick W. Siem
Anne P. Duke
Joseph L. Weeke

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Armstrong Laboratory
Human Resources Directorate
Manpower and Personnel Research Division
7909 Lindbergh Drive
Brooks Air Force Base, Texas 78235-5352

8. PERFORMING ORGANIZATION REPORT NUMBER

AL/HR-TP-1995-0008

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

Previously published in "Proceedings: Applied Behavioual Sciences Symposium," USAF Academy, 1994.
Armstrong Laboratory Technical Monitor: Dr. Frederick Siem, (210) 536-3956

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

A study was conducted to determine the dimensionality of United States Air Force pilot combat performance using performance incidents from Desert Shield/Storm. Subjects were 265 operational pilots representing seven different aircraft platforms. Some pilots generated critical incidents representing superior and average combat performance, and others sorted the incidents into categories corresponding to similar behaviors. The co-occurrence of incidents in the same performance category was used as proximity data for multidimensional scaling analyses. A combination of data analyses and expert judgment indicated that pilot combat performance could be defined by the following six dimensions: (a) Compliance with Regulations, (b) Knowledge, Skill, and Ability, (c) Crew Management, (d) Leadership, (e) Situational Awareness, and (f) Planning. Use of these performance dimensions for measurement of effective performance and for pilot selection test development is described.

14. SUBJECT TERMS

Combat Performance
Dimensionality
Situational awareness

15. NUMBER OF PAGES

8

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNLIMITED

Dimensions of Air Force Pilot Combat Performance

Introduction

The continuous search to improve military pilot selection procedures has compelled many researchers to focus attention on individual differences in human attributes as predictors of pilot performance (e.g., Carretta, 1990; Croll, Mullins, & Weeks, 1973). Proper interpretation of pilot selection research requires a suitable framework for conceptualizing the dimensions of combat performance. Validating models of pilot performance requires attention both to measures of individual differences, such as aptitude and personality, and to measures of flying performance. Most validation research has concentrated on predicting early pilot training performance (Carretta & Ree, in press).

Other than during World War II (WWII), there have been few attempts to analyze combat performance. A review of the literature identified several characteristics related to effective pilot performance. Jenkins, Ewart, and Carroll (1950) examined peer ratings from 2,872 combat pilots and identified the following characteristics associated with higher ratings of combat effectiveness: leadership/responsibility, teamwork, practical intelligence, combat aggressiveness, skill/interest in flying, conscientiousness, steadiness, and sociability. Bair (1952) performed a qualitative analysis of data describing best and worst cadets known by WWII combat Navy pilots. The characteristics found were teamwork/consideration for others, desire to fly/flying skill, personal stability/calmness, social adaptability/easy-going temperament, and conscientiousness/ability to accept responsibility.

More recent research contributes to the understanding of job performance dimensions, both for jobs in general and specifically for pilots in crew aircraft. Campbell (1990) has proposed a taxonomy of major performance components that acknowledges the multidimensionality of job-related behavior. The eight components include: job-specific task proficiency, non-job-specific task proficiency, maintaining personal discipline, demonstrating effort, communication, facilitating peer and team performance, supervision, and management/administration. These components are comparable to those from other models of performance, such as Helmreich and Foushee's (1993) model of flight crew performance which includes aircraft control tasks, procedural tasks, situational awareness, communications and decision tasks, and team formation and management tasks. The present research was stimulated by the belief that an exploratory analysis of operational and combat experiences during Desert Shield/Storm would provide the most current snapshot of pilot combat performance in the context of present combat operations doctrine and the latest weapon system technology.

Method

Participants

The participants were 265 Air Force pilots. The sample consisted of instructor pilots, co-pilots, and pilots who were assigned and on current flying status with one of seven aircraft weapon systems: bombers (B-52), fighters (F-15, F-16, F-111), transports (C-141, C-130), and special operations (AC/HC/MC-130). The majority of these pilots were captains with a minimum of six years in service. Many of the pilots had combat flying experience in Desert Shield/Storm (n = 138).

Procedure

Data collection took place at seven Air Force bases. On each data collection trip, the research team randomly assigned pilots to one of two groups (Group I, Group II). The team informed Group I (n = 91) of the purpose and goals of the research project and instructed them on how to write a critical incident according to methods outlined by Smith and Kendall (1963). The instruction included an emphasis on writing incidents involving combat experience. However, if a pilot believed a non-combat incident more clearly illustrated the difference between the exceptional and average pilot, the incident could be included. The format of a critical incident (Bownas & Bernardin, 1988) included a very brief background which established the scenario, followed by one specific observable behavior, and an immediate outcome or consequence of that behavior.

Each pilot in Group II (n = 49) independently read the incidents generated by Group I. The task for subjects in Group II was to sort incidents into categories where the incidents in a category were more similar to each other than to incidents in any other category. Constraints on the sort were that each pilot had to have greater than or equal to two categories of incidents but less than or equal to 15 categories. The pilots in Group II did not receive any predetermined category names in which to sort incidents because the purpose was to discover the dimensions underlying performance without the influence of experimenter effects. The pilots were instructed to focus on the behavior portion of each incident rather than on the resultant outcome or consequence to avoid sorts into only good and bad categories. After completing this sorting task, the Group II pilots provided labels for each category. The sorting data were used to generate inter-incident co-occurrence associations to use as proximity data for multidimensional scaling (MDS) analyses (Rosenberg & Kim, 1975).

The reliability of the sorting results was assessed during a subsequent data collection trip. Group III (n = 125) was selected to re-perform the sorting task (i.e., retranslate). None of the subjects in Group III had participated in the previous exercises. Group III's task was to read incidents for their aircraft and then to sort the incidents into named categories where the category names issued from analyses of Group II's sorting decisions.

Analyses

Constraints on analyses were to discover structure underlying combat performance while maintaining an empirical anchor in the form of pilots' actual observations of combat performance. The first analytical approach was MDS. This approach is like exploratory factor analysis in that it is a descriptive statistical technique used to determine data structure based on associations.

For each pilot in Group II, an m by m matrix of inter-incident associations was generated, with m being the number of incidents sorted into categories for a particular weapon system. The entries in the matrix represented the frequency with which any two incidents were sorted into the same category. These data were accumulated across pilots for each of the seven weapon systems using the accumulation rule for co-occurrence proximities defined by Rosenberg and Kim (1975). Each of seven matrices representing one of seven aircraft platforms was then analyzed using Alternating Least Squares Scaling (ALSCAL; Young & Lewyckyj, 1979) MDS analysis. The objective was to locate all incidents from a given platform in multidimensional space and from this geometric representation to identify the minimum number of dimensions that account for the observed data structure.

The second analytical approach was to conduct an analysis of the category labels provided by each pilot. This amounted to conducting an analysis of the sorting data at a higher level of generality than at the incident level. This analytical approach was based on the dimensional coordinates obtained from the MDS analysis of the incident data. The objective was to locate the category labels provided by each pilot along each of the six dimensions. The assumptions of this analysis were: first, the dimensional coordinates that best represented a pilot's category label were the average of the coordinates for all incidents in that category; second, the structure underlying combat performance consisted at a minimum of six dimensions. This second assumption was adopted because of the exploratory nature of the study and to avoid imposing an arbitrary ceiling on the number of dimensions in addition to that imposed by the MDS program.

This second analysis was conducted for each platform. For each pilot there were available from 2 to 15 categories with a label for each one. Different pilots had different numbers of incidents in each category, different number of categories, and therefore a different number of category labels. For each platform and each dimension, to locate the category labels for all pilots on the first dimension, the labels were assigned a coordinate value equal to the average of the coordinate values of the incidents in the category with which it was associated. After label coordinate values for all pilots were determined, the values were ranked from high to low, and labels at the extremes of the ranking were evaluated to ascertain the meaning of the dimension. This procedure was followed for each of the six dimensions so that label coordinate values for all pilots were ranked six times, once for each dimension using coordinate values from that dimension.

Results

Table 1 lists the number of pilots in Group I, the total number of incidents they produced, the number of pilots in Group II, and the average number of categories produced from the sorting exercise.

Table 1
Critical Incident Production and Categories

Weapon System	N Group I	N Incidents	N Group II	Average N Categories
AC/HC/MC-130	12	143	6	7.2
B-52	11	110	11	8.5
F/EF-111	11	100	7	7.1
C-141	13	122	8	7.8
F-16	17	163	6	6.8
C-130	13	80	6	7.8
F-15	14	116	5	5.0

At the incident level of analysis, the ALSCAL solutions for some weapon systems (C-141, C-130, and F-16) appeared to suggest two underlying dimensions, whereas the solutions for other aircraft systems suggested only one dimension. For all platforms, the one dimensional solution yielded R^2 s greater than or equal to .85. For several platforms, examination of the location of individual incidents at the extremes of the first dimension indicated effective behaviors at one extreme and incidents with ineffective behaviors at the other extreme. Evidently, despite precautions to avoid a good versus bad behavior category, this structure resided in the inter-incident associations and masked the existence of underlying dimensions.

At the label level of analysis, category labels were identified at the extremes of each of six dimensions by specific weapon systems. Representative category labels include high knowledge and ability in flight versus procedural errors, ability to prioritize versus no situational awareness, working with people versus poor communication, takes charge versus quits doing the job, poor mission preparation versus prepares for all contingencies and adherence to directives versus breaking the rules. Inspection of the labels at the extremes suggested several dimensions common across aircraft and informed the subsequent content analysis.

The content analysis of the category labels suggested eight performance dimensions that were common across all seven weapon systems. Due to significant overlap with other performance categories, two of these categories (Personal/Interpersonal Factors and Decision Making) were not applied by Group III during their re-sort task. The two omitted categories appeared to belong at a higher, more general level of classification and could be described as having a function in almost every one of the remaining six performance categories. The resulting pilot performance dimensions were as follows: (a) Compliance with Regulations (compliance or noncompliance), (b) Knowledge, Skill, and Ability (flying skills and knowledge), (c) Crew Management (crew management and utilization/mutual support), (d) Leadership, (e) Situational Awareness, and (f) Planning.

Analysis of the retranslation data indicated that on the average 69 percent of the critical incidents were sorted into one of the six performance dimensions, with 61 percent being the minimum (C-141s) and 73 percent being the maximum (F-15s).

Discussion

The results from this study replicate earlier research findings (i.e., Bair, 1952; Jenkins et al., 1950) and also provide empirical support for conceptual models of pilot performance previously mentioned. The six performance categories identified in the present study are consistent with the taxonomy of job performance discussed by Campbell (1990) and with Helmreich and Foushee's (1993) flight crew performance model. Moreover, the results of the present study suggest that the Helmreich and Foushee (1993) model extends to military aircrews in both crew and single-seat aircraft.

In terms of applying the results of the present study to performance measurement, the data suggest that pilots should be evaluated on six dimensions. For this purpose, the incidents collected for this study can serve as material for the development of pilot performance rating scales. To enhance the value of such scales, data are being collected on the effectiveness level of the behavior from each incident that was successfully retranslated into one of the six performance categories. These effectiveness level ratings can then be used to produce Behaviorally Anchored Rating Scales (BARS; Bownas & Bernardin, 1988) which also can be used in test validation research.

Finally, the results of the present study have implications for pilot selection test research. Examination of the content of the present, automated test battery specifically used for pilot selection (Carretta, 1990) indicates that it measures abilities that underlie flying skills and situational awareness and currently does not measure attributes that would underlie leadership and crew management. This emphasizes the importance of studying the validity of interpersonal behavioral skills and personality measures as pilot selection factors in future research.

References

Bair, J. T. (1952). The characteristics of the wanted and unwanted pilot in training and in combat (Memorandum Rep. No. 2). Pensacola, FL: U. S. School of Naval Aviation Medicine.

Bownas, D. A., & Bernardin, H. J. (1988). Critical incident technique. In S. Gael (Ed.), The job analysis handbook for business, industry, and government (Vol. 2). New York: John Wiley & Sons.

Campbell, J. P. (1990). Modeling the performance prediction problem in industrial and organizational psychology. In M. D. Dunnette & L. M. Hough (Eds.), Handbook of industrial and organizational psychology (Vol. 1; 2nd ed.; pp. 687-732). Palo Alto, CA: Consulting Psychologists Press, Inc.

Carretta, T. R. (1990). Cross-validation of experimental USAF pilot training performance models. Military Psychology, 2(4), 257-264.

Carretta, T. R., & Ree, M. J. (in press). Air Force officer qualifying test validity for predicting pilot training performance. Journal of Business and Psychology.

Croll, P. R., Mullins, C. J., & Weeks, J. L. (1973). Validation of the cross-cultural aircrew aptitude battery on a Vietnamese pilot trainee sample. (AFHRL-TR-73-30, AD-778 072). Lackland AFB, TX: Air Force Human Resources Laboratory, Personnel Research Division.

Helmreich, R. L., & Foushee, H. C. (1993). Why crew resource management? Empirical and theoretical bases of human factors training in aviation. In E. L. Wiener, B. G. Kanki, & R. L. Helmreich (Eds.), Cockpit resource management (pp. 3-45). San Diego, CA: Academic Press, Inc.

Jenkins, J. G., Ewart, E. S., & Carroll, J. B. (1950). The combat criterion in naval aviation (Rep. No. 6). Washington, DC: National Research Council Committee on Aviation Psychology.

Rosenberg, S., & Kim, M. P. (1975). The method of sorting as a data-gathering procedure in multivariate research. Multivariate Behavioral Research, 10, 489-502.

Smith, P. C., & Kendall, L. M. (1963). Retranslation of expectations: An approach to the construction of unambiguous anchors for rating scales. Journal of Applied Psychology, 47, 149-155.

Young, F. W., & Lewyckyj, R. (1979). ALSCAL-4 user's guide. Chapel Hill, NC: Young Psychometric Laboratories.