AL-TP-1991-0022




# THE DEVELOPMENT OF BEHAVIORALLY ANCHORED RATING SCALES (BARS) FOR EVALUATING USAF PILOT TRAINING PERFORMANCE

Thomas R. Carretta
Laurie C. Walters

HUMAN RESOURCES DIRECTORATE
MANPOWER AND PERSONNEL RESEARCH DIVISION
Brooks Air Force Base, TX 78235-5000



July 1991

Interim Technical Paper for Period October 1989 – March 1991





AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235-5000

ARMSTRONG LABORATORY

## NOTICES

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.

Thomas R. Carretta

THOMAS R. CARRETTA
Project Scientist

WILLIAM E. ALLEY, Technical Director
Manpower and Personnel Research Division

ROGER W. ALFORD, Lt Col, 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 |
|---|---|---|
| | July 1991 | Interim Paper – October 1989 – March 1991 |

**4. TITLE AND SUBTITLE**
The Development of Behaviorally Anchored Rating Scales (BARS) for Evaluating USAF Pilot Training Performance

**5. FUNDING NUMBERS**
PE - 62703F
PR - 7719
TA - 18
WU - 45

**6. AUTHOR(S)**
Thomas R. Carretta
Laurie C. Walters

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
Armstrong Laboratory
Human Resources Directorate
Manpower and Personnel Research Division
Brooks Air Force Base, TX 78235-5000

**8. PERFORMING ORGANIZATION REPORT NUMBER**
AL-TP-1991-0022

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

**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**

**11. SUPPLEMENTARY NOTES**

**12a. DISTRIBUTION/AVAILABILITY STATEMENT**
Approved for public release; distribution is unlimited.

**12b. DISTRIBUTION CODE**

**13. ABSTRACT** (Maximum 200 words)

The purpose of this study was to develop Behaviorally Anchored Rating Scales (BARS) which could be used by instructor pilots (IPs) to evaluate their students on eight personality characteristics considered important to flying fighter-type aircraft. IPs generated behavioral examples which reflected good, average, and poor job behaviors for each personality dimension. These job behaviors were randomized and presented to another group of IPs who tried to "match" each behavior with the personality characteristic it best represented. The IPs demonstrated sufficient agreement to develop BARS for four of the eight personality characteristics (achievement motivation, assertiveness, cooperativeness, and stress tolerance). The behavioral examples generated for the retained personality characteristics were evaluated for their use as scale anchor points. Several uses of BARS in the flying training environment were discussed.

**14. SUBJECT TERMS**
personality assessment
rating scales
pilot characteristics

**15. NUMBER OF PAGES**
36

**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 | UL |

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

CONTENTS

# LIST OF TABLES

## PREFACE

This project was conducted under work unit 77191845 in support of Request for Personnel Research 78-11, Selection for Undergraduate Pilot Training, issued by Air Training Command. It was initiated in response to an effort by the Euro-NATO Aircrew Selection Working Group (ACSWG) to develop selection instruments and training performance criteria for pilot candidates attending Euro-NATO Joint Jet Pilot Training (ENJJPT) at Sheppard Air Force Base, TX. Similar projects are being conducted by NATO Air Force representatives who attend the ACSWG meetings.

Appreciation is extended to Lt Col Rick Perry for his help recruiting USAF T-37 instructor pilots who served as subject matter experts in this study and to Mr Robert Piccacio for administrative support. The authors also thank Maj Dave Perry, Drs Rick Siem, Joseph L. Weeks, William E. Alley, Mike McAnulty, and Ms Theresa A. Mercatante for their comments and technical support during this project.

# THE DEVELOPMENT OF BEHAVIORALLY ANCHORED RATING SCALES (BARS) FOR EVALUATING USAF PILOT TRAINING PERFORMANCE


## SUMMARY

The purpose of this study was to determine whether experienced USAF instructor pilots could make reliable judgements about pilot candidates' personality and whether these judgements could be used to develop performance indices for pilot candidates attending Undergraduate Pilot Training (UPT).

Eight personality characteristics identified as crucial for predicting fighter-type pilot performance (Rodgers & Sage, 1986) were targeted for development of Behaviorally Anchored Rating Scales or BARS (Smith & Kendall, 1963). The unique characteristic of this type of rating scale is that the scale values are defined by examples of job-related behaviors which reflect different amounts of effectiveness on the performance dimension under consideration.

Two groups of USAF T-37 instructor pilots participated as subject matter experts (SMEs) in this study (20 in group 1 and 22 in group 2). SMEs demonstrated sufficient agreement to develop BARS for four of the eight personality characteristics targeted in this study (achievement motivation, assertiveness, cooperativeness, and stress tolerance). SMEs often had difficulty distinguishing among the remaining characteristics (e.g., aggressiveness, decisiveness, leadership, and risk-taking examples often were cross-classified). As a result, these characteristics were eliminated from consideration for final BARS development.

The resulting scales can be used in several ways. For example, USAF instructor pilots could use scale ratings to provide pilot candidates with performance feedback (i.e., counseling, training). BARS ratings also could be used as training performance criteria for preselection personnel assessment instruments (i.e., to validate sources of personnel data used to select pilot candidates.)


## INTRODUCTION

Personnel researchers involved in the selection and classification of both military and civilian aviators have long been interested in identifying personality characteristics of successful (versus unsuccessful) pilots. The literature is voluminous, albeit inconclusive (Dolgin & Gibb, 1988). Attempts to establish a link between personality and pilot success have been somewhat disappointing, and understandably so, due to such factors as varying methodologies and performance criteria across studies. Moreover, the issues inherent in any type of personality research (e.g., predictive/explanatory ability of global vs. specific traits) certainly add to the difficulty of drawing conclusions from the literature. Although recent studies have identified personality characteristics associated with successful pilot performance (e.g., Carretta & Siem, 1988; Siem, 1990), the utility of personality tests for selection of military pilots has not been clearly established (Dolgin & Gibb, 1988; North & Griffin, 1977; Siem, 1990).

The traditional approach for investigating the impact of personality on pilot performance has been to administer some combination of psychological tests and then analyze the relationship between personality profiles and pilot performance. These self-report tests are susceptible to "faking," as subjects may select responses that seem appropriate for the job or are socially desirable, as opposed to selecting their "true" responses. Also, the actor-observer and self-serving biases as discussed in the social psychological literature (Jones & Nisbitt, 1972) suggest that individuals tend to attribute their own behavior to situational, rather than dispositional causes, except when it is "self-serving" to do otherwise.

The present study sought to develop performance-based rating scales which could be used by instructor pilots to evaluate their students on various personality characteristics considered important to flying training performance. According to Funder (1990), when conducting personality research one

should consider (a) how personality is manifested in behavior and (b) how personality can be judged accurately. The development of Behaviorally Anchored Rating Scales (BARS) (Smith & Kendall, 1963) as a tool for assessing personality and performance in pilot training addresses these two issues.

The first issue regarding how personality manifests in behavior involves expressing global traits in behavioral terms. Assuming that trait expression is at least somewhat influenced by the situation, it makes sense to define behavior corresponding to each personality construct in the context of a flying training environment where the behavior will be evaluated. Eight personality traits were identified by a group of subject matter experts (SMEs) as being relevant to flying fighter-type aircraft (see Appendix A). Another group of SMEs generated behavioral examples of each personality construct. These generated behaviors were later considered for inclusion in the BARS.

It may seem quite odd that pilots, rather than psychologists, are defining personality constructs. However, given that experienced pilots have their own implicit personality theories as to what makes a good pilot (e.g., assertiveness, achievement-orientation, etc.), it is beneficial for researchers to explicitly define the pilots' theories to better understand the information used when they evaluate their students. Hogan and Hogan (1991) note that "personality has its social impact in terms of qualities ascribed to individuals by their friends, employers, and colleagues" (p. 12). So, if an instructor pilot observes a student over a period of time and characterizes him or her as assertive, then for all intent and purposes, that student *is* assertive.

The second issue of how personality can be accurately judged is the primary impetus behind the development and application of BARS in the flying training environment. Given the interest in personality and pilot performance, it is important to define, judge, and "measure" personality as accurately as possible to reduce error in prediction models or any other type of statistical analyses. The approach taken in this study is to let experienced pilots be "judges" of personality. This research effort focuses primarily on the development of BARS to be used in a flying training environment.

The BARS approach was adopted for several reasons. A unique characteristic of this type of rating scale is that the attribute (i.e., personality construct) represented by the scale is defined by observable behaviors which reflect different amounts of effectiveness on the dimension under consideration (Smith & Kendall, 1963). This feature of the scale attempts to address the issue regarding the manifestation of personality in behaviors. Another positive aspect of using BARS is that it gets the people who ultimately will be using the scales (i.e., instructor pilots) to participate in scale development. As a result, scale values are expressed in behavioral terms that are easily understood by the people using it. Also, because the "users" of the scale participate in the development phase, there may be an increased motivation to do an effective job of evaluation, and similarly, appraisees may be more likely to accept the results (Dunnette, 1966). BARS have been used in a variety of occupational settings to evaluate nurses, police officers, counselors, and grocery clerks (for a review, see Schwab, Heneman, & DeCotiis, 1975).

The BARS format used by Smith and Kendall (1963) is intended to provide a standard meaning for the points on the rating scale and therefore improve ratings of job or training performance (Bernardin & Smith, 1981). The Smith and Kendall format and development procedure have been used widely either in the original form or with variation (Landy & Farr, 1983). Table 1 summarizes the procedural sequence used by Smith and Kendall for the development of BARS.

## METHOD

### Subjects

Three groups of experienced military pilots participated as SMEs. Table 2 provides a description of the groups and their functions in this study. All participants were informed their participation was

voluntary, their responses would be confidential, and the data would be used only to improve pilot candidate selection and classification procedures. Instructor pilots responsible for the early stages of training (i.e., T-37 instructors) were used in groups B and C because it was felt that the BARS should reflect performance in the early stages of training (i.e., before too many of the weaker pilot candidates had been eliminated).

TABLE 1. PROCEDURE FOR DEVELOPING BEHAVIORALLY ANCHORED RATING SCALES (SMITH & KENDALL, 1963)

| | |
|---|---|
| Identification and definition of performance dimensions | Group A identifies all important dimensions for job performance. They also define conceptually each performance dimension and define good, average, and poor performance on each dimension. |
| Generation of behavioral examples | Group B gives examples of good, average, and poor job behaviors for each performance dimension. |
| Retranslation and allocation | Group C is presented with a randomized list of behavioral examples and a list of performance dimensions. Each SME allocates or classifies each behavioral example to the performance dimension it best represents. |
| Scaling | Group D evaluates behavioral examples in terms of the effectiveness of the performance described. |
| Scale anchor selection | Personnel researcher computes the mean and standard deviation of the rating given to the behavioral examples. Examples are selected as anchors such that items have mean values that provide anchors for the entire performance scale (poor to good) and that items have relatively small standard deviations. |

TABLE 2. SUBJECT MATTER EXPERTS PARTICIPATING IN THE DEVELOPMENT OF BEHAVIORALLY ANCHORED RATING SCALES

| Group | N | Group Description | Function |
|---|---|---|---|
| A | 43 | Experienced fighter pilots, from Canada (n = 18), Norway (n = 15), and the United States (n = 10) Air Force | Identification and definition of performance dimensions |
| B | 20 | USAF T-37 instructor pilots | Generation of behavioral examples |
| C | 22 | USAF T-37 instructor pilots | Retranslation and allocation; scaling |

## Procedure

The procedure used was an iterative one. Products created by one group were checked, revised, and augmented by successive groups.

Group A performed an analysis to determine the importance of personnel characteristics for fighter-type pilot tasks using a modified version of the Levine, Mallamad, and Fleishman (1978) decision-flow diagram. In this technique, SMEs use a structured questionnaire to answer a series of questions to determine the importance of characteristics for performing 12 job-related tasks. Results of this phase specified the relative importance of 27 personnel characteristics when selecting fighter-type pilot candidates. Eight of the 27 characteristics were personality dimensions considered relevant to flying fighter-type aircraft. Definitions of the personality characteristics are provided in Appendix A and the 12 pilot tasks are summarized in Appendix B.

Group A SMEs were provided with a package containing the decision-flow diagram, 12 critical pilot tasks, and a set of instructions. Each SME was required to select a task (e.g., perform tactical defensive flight maneuvers), and while considering the task, work his/her way through the decision-flow diagram. Each time the SME decided a personal characteristic was necessary for successful completion of the task under consideration, he/she indicated its importance on a five-point scale. Once all questions in the decision-flow diagram had been completed for a particular task, a new task was selected and the process repeated. Using this procedure, the SMEs independently rated the importance of each of the 27 characteristics for the performance of each of 12 critical pilot tasks.

Group B generated behavioral examples of good and poor performance during pilot training for each of the eight personality characteristics identified by Group A. Due to the large number of unique behavioral examples (n = 261) generated by Group B, it was necessary to reduce the list of behaviors prior to the retranslation/allocation phase. A team consisting of a personnel research psychologist and two SMEs combined similar examples and deleted others that were inappropriate (i.e., not examples of flying training performance), unclear, or poorly defined.

Group C performed both the retranslation/allocation phase and scaling phase. Each SME was presented with definitions of the eight personality characteristics and a deck of 73 index cards. Each card had one behavioral example printed on it. The SMEs were instructed to familiarize themselves with the definitions of the eight personality characteristics, then read the behavioral examples on each of the 73 index cards and place each example under the personality characteristic to which it was related most closely. When each SME had completed sorting the index cards (i.e., completed retranslation/allocation), he/she went back to each pile and indicated whether each behavior reflected (1) poor (2) below average (3) above average, or (4) excellent performance (i.e., rated the behavioral examples for each personality characteristic).

Personnel research psychologists reviewed results from Group C to identify behavioral anchors for each point on the rating scale for each personality characteristic. In order to be useful as scale anchors, the Group C SMEs must agree on assignment of behavioral examples to the personality characteristics and on the magnitude of the personality characteristics represented by the behavioral example.

## Scale Format Issues

To facilitate the use of BARs, Smith and Kendall (1963) recommend that the behavioral anchors be worded in a "could be expected to" format (e.g., "The student could be expected to make excuses for his/her performance.") By using the expectation format, the rater is not being asked to indicate whether or not the ratee had been observed performing a particular behavior. Instead, the rater is being asked to infer or predict the ratee's behavior in terms of the scale anchors based on the rater's prior observations of the ratee's job performance. For a given ratee, it is quite likely that the rater will not have had the opportunity to observe whether the ratee performed in a manner described by the

behavioral anchors. This problem is avoided by asking the rater to make a prediction about the ratee's expected training performance. To clarify the basis of the rater's evaluation, Landy and Farr (1983) recommend requiring written statements from the rater that describe behavioral examples that substantiate the rating for a given personality factor.

Another problem that may occur when using BARS to assess training performance is the objection of raters to the order of behavioral anchors on a given scale. Although behavioral anchors are chosen that demonstrated acceptable agreement among SMEs during the scaling phase (i.e., small standard deviations for those ratings), it is common for the confidence intervals around the means of adjacent behavioral anchors to overlap. As a result, some raters may disagree with the position of the behavioral anchor on the scale. Raters may question the credibility of the BARS when this occurs. Landy and Farr (1983) suggest that grouping behavioral examples into broad performance-level categories (i.e., more than adequate, adequate, and less than adequate) may help to minimize this problem.

The scales presented in Appendix E incorporate several of these refinements. The instructions for these scales stress that the behavioral anchors represent a sample of possible training behaviors that may or may not have been observed for a given ratee. Also the scale anchors are grouped into three broad categories (i.e., more than adequate, adequate, and less than adequate) and the rater is required to provide a specific example of the training behavior to substantiate the rating.

## RESULTS

Given the iterative nature of the scale development, it is important to consider the results obtained from each procedural phase. Therefore, the following results are organized in accordance with the developmental sequence outlined by Smith and Kendall (1963).

### Identification and Definition of Performance Dimensions

An ability analysis of fighter-type pilot tasks using a modified version of the Levine et al. (1978) decision flow diagram identified eight personality characteristics critical to job performance (Rodgers & Sage, 1986). The eight personality characteristics included achievement motivation, aggressiveness, assertiveness, cooperativeness, decisiveness, leadership, risk taking, and stress tolerance. Definitions of these characteristics are provided in Appendix A and the critical fighter-type pilot tasks are summarized in Appendix B.

### Generation of Behavioral Examples

Two hundred sixty-one (261) unique behavioral examples were generated by the 20 SMEs who participated in this phase. A personnel research psychologist and two different SMEs reviewed these examples with the goal of reducing them to a more manageable number for the retranslation and allocation phase. Some similar items were combined while others were eliminated because they were not clear or did not appear to reflect the attribute for which they were generated, or were not examples of training performance. This process reduced the number of examples from 261 to 73. Table 3 summarizes the number of unique behavioral examples grouped by personality characteristic that were generated and retained for the retranslation and allocation phase. Appendix C lists the 73 behavioral examples that were retained for the retranslation and allocation phase.

### Retranslation and Allocation

Criteria for successful retranslation and allocation vary among researchers describing the BARS process. Smith and Kendall (1963) recommend that behavioral examples be retained only if (a) there is clear model agreement among SMEs as to the characteristic to which the behavioral example belongs and (b) the example is allocated to the same characteristic for which it was designed during the behavior generation phase. Landy and Farr (1983) suggest that clear agreement among SMEs (e.g., 70%

consensus) is sufficient to retain the behavioral example for the scaling phase. The Landy and Farr (1983) method was used in this study. Examples were judged to be retranslated/allocated accurately if 15 of the 22 SMEs (68.2%) placed it under the same characteristic. This approach was adopted because it appeared that the SMEs found it difficult to distinguish among some of the characteristics (e.g., aggressiveness and decisiveness).

TABLE 3. NUMBER OF BEHAVIORAL EXAMPLES ORIGINALLY GENERATED AND RETAINED FOR THE RETRANSLATION AND ALLOCATION PHASE

| Characteristic | Number of Behaviors Originally Generated | Number of Behaviors Retained |
|---|---|---|
| Achievement Motivation | 32 | 11 |
| Cooperativeness | 21 | 8 |
| Stress Tolerance | 43 | 9 |
| Assertiveness | 30 | 8 |
| Aggressiveness | 35 | 11 |
| Risk Taking | 24 | 4 |
| Leadership | 29 | 8 |
| Decisiveness | 47 | 14 |
| Totals | 261 | 73 |

Results from the retranslation and allocation phase are summarized in Table 4. SMEs were able to reach acceptable (i.e., 70%) agreement for 50.7% (37 of 73) of the behavioral examples. It should be noted that out of the 37 examples, only 26 were allocated to the same characteristic for which the example was generated. The other 11 of 37 behavioral examples where sufficient agreement was obtained were cross-translated (i.e., behavioral examples that were allocated with 70% agreement to a different characteristic than the one for which they originally were generated). Agreement was strongest for the characteristics of achievement motivation, cooperativeness, stress tolerance, and assertiveness. An examination of the distribution of the behavioral examples suggests that SMEs often had trouble distinguishing among several characteristics. Behavioral examples for aggressiveness, decisiveness, leadership, and risk-taking were often cross-classified (i.e., the same example was assigned under two or more characteristics). As a result, these characteristics were eliminated from consideration for final BARS development.

## Scaling

Scale values (i.e., means and standard deviations) which quantified the magnitudes of a given personality characteristic represented by the behavioral examples were examined for the four characteristics that demonstrated acceptable agreement in the retranslation/allocation phase (achievement motivation, assertiveness, cooperativeness, and stress tolerance) and are presented in

Appendix D. SMEs demonstrated sufficient agreement as to the location of the behavioral examples on these scales.

TABLE 4. NUMBER OF BEHAVIORAL EXAMPLES RETRANSLATED[a] AND CROSS-TRANSLATED[b] FOR EACH PERSONALITY CHARACTERISTIC

| Characteristic | Number of Behaviors | Number of Behaviors Retranslated From Same Category | Number of Behaviors Cross-Translated From Other Categories | Total Number of Behaviors Retained |
|---|---|---|---|---|
| Achievement Motivation | 11 | 7 | 1 | 8 |
| Cooperativeness | 8 | 5 | 2 | 7 |
| Stress Tolerance | 9 | 4 | 1 | 5 |
| Assertiveness | 8 | 4 | 1 | 5 |
| Aggressiveness | 11 | 1 | 1 | 2 |
| Risk Taking | 4 | 1 | 2 | 3 |
| Leadership | 8 | 2 | 1 | 3 |
| Decisiveness | 14 | 2 | 2 | 4 |
| Totals | 73 | 26 | 11 | 37 |

[a]The term "Retranslated" refers to behavioral examples that were successfully allocated to the same characteristic under which they were originally generated.

[b]The term "Cross-Translated" refers to behavioral examples that were allocated to a different characteristic (70% agreement among SMEs) than the one under which they originally were generated.

## DISCUSSION

### Inadequately Defined Scales

Sufficient agreement was achieved among SMEs as to the behaviors associated with different characteristics to develop BARS for four of eight target characteristics (i.e., achievement motivation, assertiveness, cooperativeness, and stress tolerance). This does not imply that these characteristics are more important or more closely related to fighter-type pilot performance than the four characteristics for which BARS were not developed (i.e., aggressiveness, decisiveness, leadership, and risk-taking). Examination of the results of the retranslation and allocation phase and discussions with other SMEs suggest that sufficient agreement was not achieved for these characteristics because (a) the SMEs had difficulty making meaningful distinctions among the definitions of these characteristics (i.e., the

characteristics are related closely to each other) and (b) many of the behavioral examples generated are related to more than one characteristic. For example, the statement "pilot candidate is not afraid to make decisions based on incomplete information" could be interpreted as an indication of several characteristics (e.g., aggressiveness, assertiveness, decisiveness, leadership, or risk-taking) depending upon its context (i.e., the characteristic being stressed in the situation, the appropriateness of the behavior).

A follow-up study could be performed to determine the extent to which some characteristics and their behavioral examples are perceived as similar to other characteristics and their examples. For example, during the retranslation and allocation phase, SMEs could be instructed to place each behavioral example under as many characteristics as he/she feels it is related. A factor analytic approach could be used to determine the extent to which the different characteristics are related to each other.

**Adequately Defined Scales**

BARS for the four characteristics that were adequately defined are presented in Appendix E (achievement motivation, assertiveness, cooperativeness, and stress tolerance). These scales can be used by instructor pilots to provide performance feedback to pilot candidates (i.e., counseling, training) or as training performance criteria against which preselection personnel assessment instruments could be validated.

As previously discussed, the purpose of this study was to develop BARS for eight critical fighter-type personality characteristics identified by Rodgers and Sage (1986). Future research should consider the relationship between personality as defined and measured by BARS, with personality as measured by individual self-report inventories, and by peer ratings. The BARS approach also could be expanded to develop scales for the nonpersonality performance dimensions discussed by Rodgers and Sage (e.g., memorization, psychomotor coordination, reasoning, and situational awareness).

## CONCLUSION

USAF instructor pilots were able to reach sufficient agreement (as to the behaviors associated with personality characteristics) to develop BARS for four of eight target characteristics. The resulting scales can be used in several ways. For example, USAF instructor pilots could use scale ratings as a means to provide pilot candidates with performance feedback (i.e., counseling, training). BARS also could be used as training performance criteria for developing preselection personnel assessment instruments (i.e., to validate sources of personnel attribute data used to select pilot candidates).

Follow on studies are planned to develop BARs for nonpersonality pilot characteristics identified by Rodgers and Sage (1986) and to evaluate these scales in an operational environment.

## REFERENCES


Bernardin, H.J., & Smith, P.C. (1981). A clarification of some issues regarding the development and use of behaviorally anchored rating scales (BARS). *Journal of Applied Psychology, 66, 4*, 458-463.

Carretta, T.R., & Siem, F.M. (1988). *Personality, attitudes and pilot training performance: Final analyses* (AFHRL-TP-88-23, AD-A199 983). Brooks Air Force Base, TX: Air Force Human Resources Laboratory, Manpower and Personnel Division.

Dolgin, D.L., & Gibb, G. D. (1988). Personality assessment in aviator selection. In R.S. Jensen (Ed.), *Aviation Psychology*, Gower Publishing Company, Brooksfield, VT.

Dunnette, M.D. (1966). *Personnel selection and placement.* Belmont: Wadsworth.

Funder, D.C. (1990). Global traits: A neo-Allportian approach to personality. *Psychological Science, 2*, 31-38.

Hogan, R., & Hogan, J. (in press). Personality and status. In D.G. Gilbert & J.J. Conley (Eds.), *Personality, social skills, and psychopathology: An individual differences approach.* New York: Plenum.

Jones, E.E., & Nisbett, R.E. (1972). The actor and the observer: Divergent perceptions of the cause of behavior. In E.E. Jones, D. Kanouse, H.H. Kelley, R.E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behaviors.* Morristown, N.J.: General Learning Press.

Landy, F.J., & Farr, J.L. (1983). *The measurement of work performance: Method, theory and application.* Academic Press, Inc., Orlando, FL, 60-64.

Levine, N.G., Mallamad, M.S., & Fleishman, A.E. (1978). *Decision aids in estimating personnel requirements.* Washington, D.C.: Advanced Research Resources Organization.

North, R.A., & Griffin, G.R. (1977). *Aviator selection 1919-1977.* Special Report 77-2, Pensacola NAS, FL: Naval Aerospace Medical Research Laboratory.

Rodgers, M.N., & Sage, P.M. (1986). *Identification of ability requirements for NATO fast-jet pilot selection.* Technical Note 34/86. Willowdate, Ontario: Canadian Forces Applied Research Unit.

Schwab, D.P., Heneman, H.G., & DeCotiis, T.A. (1975). Behaviorally anchored rating scales: A review of the literature. *Personnel Psychology, 28*, 549-562.

Siem, F.M. (1990). *Predictive validity of an automated personality inventory for air force pilot selection.* (AFHRL-TP-90-55, AD-A226 731). Brooks AFB, TX: Air Force Human Resources Laboratory, Manpower and Personnel Division.

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, 2*, 149-155.

## APPENDIX A: EIGHT PERSONALITY CHARACTERISTICS CONSIDERED RELEVANT TO FLYING FIGHTER-TYPE AIRCRAFT

1. *Achievement Motivation* - a willingness and determination to work towards goals

2. *Aggressiveness* - a tendency to accept challenges, to make daring decisions, to take initiative, and to carry out actions in order to surmount difficulties

3. *Stress Tolerance* - the ability to cope with stressful circumstances (in both the short and long-term) without significant degradation in performance

4. *Risk Taking* - the willingness to make timely decisions based on limited information, while accepting responsibility for those decisions regardless of their consequences (positive or negative)

5. *Cooperativeness* - the willingness to coordinate one's own activities with other people or groups (e.g., crew members)

6. *Assertiveness* - self-assurance (i.e., belief in one's own capabilities) and willingness to defend one's own opinions

7. *Leadership* - the ability to organize the tasks of others, determine priorities, and motivate coworkers.

8. *Decisiveness* - the ability to quickly choose a course of action and to carry it out without delay

## APPENDIX B: CRITICAL FIGHTER-TYPE PILOT TASKS

1. Perform systems/weapons checks.
2. Manage on-board systems.
3. Set up attack.
4. Perform tactical offensive flight maneuvers.
5. Avoid, evade or suppress threats.
6. Monitor and control flight parameters.
7. Perform weapons delivery.
8. Perform formation tactics.
9. Respond to aircraft emergency situations.
10. Manage communications.
11. Perform low level navigation.
12. Perform tactical defensive flight maneuvers.

APPENDIX C: BEHAVIORAL EXAMPLES RETAINED FOR THE RETRANSLATION/ALLOCATION PHASE

1. *Achievement Motivation* - a willingness and determination to work towards goals.

Student is fully prepared for each mission (e.g., reviews readings required for performance of maneuvers for that flight; focuses on the task or assignment; prepared to ask questions about material that is not clear).

Student has a well-developed mission profile plan; reviews checklist.

Student participates in group study sessions.

Student tries to learn from his/her mistakes (e.g., wants to repeat a maneuver when he/she fails on first attempt).

Student is attentive during mission debriefings (e.g., takes notes).

Student shows willingness to make personal sacrifices to achieve goals (e..g.,extra study time).

Student continues to fly aircraft and listen to instructions after an obvious mistake that will ensure an unsatisfactory flight grade.

Student is not prepared for missions (e.g., does not do required readings).

Student is easily discouraged; gives up easily if he/she does not understand something.

Student lacks initiatie; student expects to be "spoon fed" information by instructors.

Student is unexcited about training; does not seem to care about his/her performance (e.g., student is apathetic during mission debriefings; shows up late for briefings or flights).

2. *Aggressiveness* - a tendency to accept challenges, to make daring decisions, to take initiative, and to carry out actions in order to surmount difficulties.

Student shows willingness to accept command of the aircraft to perform mission objectives.

Student is willing to try a new maneuver with little coaxing by IP.

Student is not afraid to volunteer.

Student has a "take charge" attitude.

Student is not afraid to talk on the radio.

Student does not accept command of the aircraft in order to perform mission objectives; routinely looks to the IP for guidance.

Student lets aircraft control him/her.

Student is reluctant to try new maneuvers. He/she asks for additional instruction before performing a new maneuver.

Student is constantly unprepared for sorties.

Student is slow in making decisions.

Student is afraid to talk on the radio.

3. *Stress Tolerance* - the ability to cope with stressful circumstances (in both the short and long-term) without significant degradation in performance.

Student confidently speaks up at Emergency Procedures meetings, presenting his/her opinion in an organized manner.

Student is generally calm; difficult to fluster.

Student is able to maintain simulator and/or aircraft control when faced with unusual or unexpected situations.

Student does not get flustered when he/she fails to perform a maneuver successfully; tries again.

Student does not speak up at Emergency Procedure meetings or speaks softly. He/she is unorganized in his/her approach to handling emergency procedures; nervous in "stand up" situation.

Student is excited easily.

When the student is faced with unexpected or unusual conditions, the instructor must intervene to keep the aircraft out of an unsafe condition.

Student gets flustered when he/she performs a maneuver poorly; then gives up and hesitates to try again.

Student "freezes" at the controls during time critical situations.

4. *Risk Taking* - the willingness to make timely decisions based on limited information, while accepting responsibility for those decisions regardless of their consequences (positive or negative).

Student is willing and eager to try new maneuvers; not concerned if IP needs to take control of the aircraft.

Student attempts to "show off"; performs dangerous maneuvers.

Student waits for IP to tell him/her to do everything; does not take initiative.

Student "dives into" a solution for a problem without thinking it through.

5. *Cooperativeness* - the willingness to coordinate one's own activities with other people or groups (e.g., crew members).

Student demonstrates leadership potential (e.g., organizes study groups, helps other students to learn material).

Student performs in-flight checks - solicits responses from others and offers own responses.

Student is a "team player" (e.g., student is willing to exchange duties/take off times with a classmate if this will help the classmate or class as a whole).

Student actively participates in class activities (on flight line and socially).

Student isolates himself/herself from others (e.g., studies alone, sits off in a corner away from others).

Student argues with IP.

Student does not participate in class activities (on flight line and socially).

Student does not help other students who are having problems with material.

6. *Assertiveness* - self-assurance (i.e., belief in oe's own capabilities) and willingness to defend one's own opinions

Student is very self-confident, not easily swayed; "take charge" attitude.

Student makes airborne decisions without IP's direction.

Student challenges IP's techniques based on book knowledge or another IP's instructions.

Student asks questions/volunteers to answer IP questions.

Student is willing to offer his/her opinion or propose solutions for a problem.

Student often needs IP's intervention.

Student is easily swayed; does not believe in his/her own ability.

Student is quiet; does not speak out when a classmate says something that is obviously wrong.

7. *Leadership* - the ability to organize the tasks of others, determine priorties, and motivate coworkers.

Student has a "take charge" attitude (e.g., makes decisions; takes initiative to organize study groups; leads discussions; positive attitude).

Student is supportive (e.g., helps others, encourages and motivates classmates).

Student volunteers for leadership positions.

Student hosts social activities.

Student demonstrates negative/cynical attitude.

Student is selfish, self-centered (e.g., when flying lead in formation, more concerned with own flying than being smooth lead).

Student waits for others to take charge; he/she is reluctant to accept responsibility; allows others or IP to make decisions - slow to act; must be told what to do.

Student is a loner; disinterested in group.

8. *Decisiveness* - the ability to quickly choose a course of action and to carry it out without delay.

Student demonstrates the ability to filter information from radio and take proper action.

Student takes charge, makes the decisions in the aircraft without asking the IP what to do.

Student reacts quickly to emergency procedure situations and is able to identify problems and initiate corrective actions.

Student is willing to make a decision, even on limited knowledge rather than doing nothing.

Student develops and carries out flight profile plan; decides to alter the mission plan himself/herself when conditions dictate.

Student lets IP make all the decisions.

Student is completely dumbfounded if things do not follow the expected course of action.

Just like the poor risk taker, the student waits for the IP to tell him/her when to turn, when to climb, when to descend, etc.

Student is just along for the ride, unsure of what to do and wastes time and fuel.

Student is wishy-washy, can't decide what to do and delays action such that the IP has to make suggestions or take over.

Student waits for IP to direct break-out.

Student is unable or slow to develop plan, lacks confidence to carry it out.

Student will not take action on emergency procedures without full knowledge.

Student becomes confused or "task saturated" when faced with a large number of tasks.

## APPENDIX D: RESULTS FROM THE RETRANSLATION/ALLOCATION AND SCALING PHASES

### TABLE D-1. ACHIEVEMENT MOTIVATION: SCALE VALUES FOR SUCCESSFULLY RETRANSLATED EXAMPLES

| Behavioral Examples | Mean | SD | N | Original Category |
|---|---|---|---|---|
| Student is fully prepared for each mission (e.g., reviews readings required for performance of maneuvers for that flight; focuses on the task or assignment; prepared to ask questions about material that is not clear) | 3.71 | .17 | 21 | Ach |
| Student shows willingness to make personal sacrifices to achieve goals (e.g., extra study time) | 3.52 | .25 | 21 | Ach |
| Student tries to learn from his/her mistakes (e.g., wants to repeat a maneuver when he/she fails on first attempt) | 3.39 | .24 | 18 | Ach |
| Student has well-developed mission profile plan; reviews checklist | 3.25 | .28 | 16 | Ach |
| Student is attentive during mission debriefings (e.g., takes notes) | 3.24 | .18 | 21 | Ach |
| Student is not prepared for missions (e.g., does not do required readings) | 1.23 | .18 | 22 | Ach |
| Student is unexcited about training; does not seem to care about his/her performance (e.g., student is apathetic during mission debriefings; shows up late for debriefings or flights) | 1.17 | .14 | 18 | Ach |
| Student is constantly unprepared for sorties | 1.10 | .08 | 21 | Agg |

**Notes.** 1. The column labeled "N" refers to the number of SMEs (out of 22) that placed the example in this category during the retranslation/allocation phase.

2. Ach = Achievement motivation
   Agg = Aggressiveness

TABLE D-2. AGGRESSIVENESS: SCALE VALUES FOR SUCCESSFULLY RETRANSLATED EXAMPLES

| Behavioral Examples | Mean | SD | N | Original Category |
|---|---|---|---|---|
| Student lets aircraft control him/her | 1.38 | .23 | 16 | Agg |
| Student waits for IP to tell him/her to do everything; does not take initiative | 1.38 | .23 | 14 | Risk |

**Notes**. 1. The column labeled "N" refers to the number of SMEs (out of 22) that placed the example in this category during the retranslation/allocation phase.

2. Agg = Aggressiveness
   Risk = Risk Taking

TABLE D-3. STRESS TOLERANCE: SCALE VALUES FOR SUCCESSFULLY RETRANSLATED EXAMPLES

| Behavioral Examples | Mean | SD | N | Original Category |
|---|---|---|---|---|
| Student is able to maintain simulator and/or aircraft control when faced with unusual or unexpected situations | 3.38 | .23 | 16 | Stress |
| Student is generally calm; difficult to fluster | 3.21 | .17 | 19 | Stress |
| Student is excited easily | 1.90 | .09 | 20 | Stress |
| Student becomes confused or "task-saturated" when faced with a large number of tasks | 1.81 | .15 | 16 | Dec |
| Student "freezes" at the controls during time critical situations | 1.06 | .06 | 17 | Stress |

**Notes**. 1. The column labeled "N" refers to the number of SMEs (out of 22) that placed the example in this category during the retranslation/allocation phase.

2. Stress = Stress-Tolerance
   Dec = Decisiveness

TABLE D-4. RISK TAKING: SCALE VALUES FOR SUCCESSFULLY RETRANSLATED EXAMPLES

| Behavioral Examples | Mean | SD | N | Original Category |
|---|---|---|---|---|
| Student is willing and eager to try new maneuvers; not concerned if IP needs to take control of aircraft | 3.54 | .25 | 13 | Risk |
| Student is willing to make a decision; even on limited knowledge rather than doing nothing | 3.39 | .24 | 18 | Dec |
| Student will not take action on EP without full knowledge | 1.77 | .18 | 13 | Dec |

Notes. 1. The column labeled "N" refers to the number of SMEs (out of 22) that placed the example in this category during the retranslation/allocational phase.

2. Risk = Risk Taking
   Dec = Decisiveness

TABLE D-5. COOPERATIVENESS: SCALE VALUES FOR SUCCESSFULLY RETRANSLATED EXAMPLES

| Behavioral Examples | Mean | SD | N | Original Category |
|---|---|---|---|---|
| Student is a "team player" (e.g., student is willing to exchange duties/take off times with a classmate if this will help the classmate or class as a whole) | 3.64 | .23 | 22 | Coop |
| Student participates in group study sessions | 3.15 | .29 | 21 | Ach |
| Student actively participates in class activities (on flight line and socially) | 3.05 | .16 | 19 | Coop |
| Student does not participate in class activities (on flight line and socially) | 1.90 | .09 | 20 | Coop |
| Student does not help other students who are having problems with material | 1.78 | .22 | 18 | Coop |
| Student is a loner; disinterested in group | 1.68 | .22 | 19 | Lead |
| Student isolates himself/herself from others (e.g., studies alone, sits off in a corner away from others | 1.50 | .25 | 20 | Coop |

Notes. 1. The column labeled "N" refers to the number of SMEs (out of 22) that placed the example in this category during the retranslation/allocation phase.

2. Coop = Cooperativeness
   Ach = Achievement Motivation
   Lead = Leadership

TABLE D-6. ASSERTIVENESS: SCALE VALUES FOR SUCCESSFULLY RETRANSLATED EXAMPLES

| Behavioral Examples | Mean | SD | N | Original Category |
|---|---|---|---|---|
| Student confidently speaks up at EP meetings, presenting his/her opinion in an organized manner | 3.71 | .20 | 14 | Stress |
| Student asks questions/volunteers to answer IP questions | 3.21 | .17 | 14 | Assert |
| Student is not afraid to talk on the radio | 3.00 | .00 | 14 | Agg |
| Student challenges IP techniques based on book knowledge or another IP's instructions | 3.00 | .62 | 16 | Assert |
| Student is afraid to talk on the radio | 1.85 | .41 | 14 | Agg |
| Student is quiet; does not speak out when a classmate says something that is obviously wrong | 1.76 | .19 | 16 | Assert |
| Student is easily swayed; does not believe in his/her own ability | 1.45 | .54 | 20 | Assert |

**Notes.** 1. The column labeled "N" refers to the number of SMEs (out of 22) that placed the example in this category during the retranslation/allocation phase.

2. Stress = Stress Tolerance
   Assert = Assertiveness
   Agg = Aggressiveness

TABLE D-7. LEADERSHIP: SCALE VALUES FOR SUCCESSFULLY RETRANSLATED EXAMPLES

| Behavioral Examples | Mean | SD | N | Original Category |
|---|---|---|---|---|
| Student has a "take charge" attitude (e.g., makes decisions; takes initiative to organize study groups; leads discussions; positive attitude) | 3.85 | .13 | 20 | Lead |
| Student demonstrates leadership potential (e.g., organizes study groups, helps other students to learn material) | 3.59 | .24 | 22 | Coop |
| Student volunteers for leadership positions | 3.35 | .23 | 17 | Lead |

**Notes**. 1. The column labeled "N" refers to the number of SMEs (out of 22) that placed the example in this category during the retranslation/allocation phase.

2. Lead = Leadership
   Coop = Cooperativeness

TABLE D-8. DECISIVENESS: SCALE VALUES FOR SUCCESSFULLY RETRANSLATED EXAMPLES

| Behavioral Examples | Mean | SD | N | Original Category |
|---|---|---|---|---|
| Student reacts quickly to EP situations and is able to identify problems and initiate corrective action | 3.86 | .12 | 14 | Dec |
| Student makes airborne decisions without IP's direction | 3.69 | .21 | 16 | Assert |
| Student is wishy-washy, can't decide what to do and delays action such that the IP has to make suggestions or take over | 1.43 | .25 | 14 | Dec |

**Notes**. 1. The column labeled "N" refers to the number of SMEs (out of 22) that placed the example in this category during the retranslation/allocation phase.

2. Dec = Decisiveness
   Assert = Assertiveness

## APPENDIX E: FINAL BEHAVIORALLY ANCHORED RATING SCALES

**Instructions for Use of Behaviorally Anchored Rating Scales**

The rating scales on the following pages were designed for the purpose of rating personality factors underlying training performance of USAF pilot candidates during T-37 training. The personality factors include (1) achievement motivation, (2) assertiveness, (3) cooperativeness, and (4) stress tolerance.

For each scale, read the definition of the performance characteristic at the top of the page and review the behavioral examples. *THE BEHAVIORAL EXAMPLES ARE ONLY A SAMPLE OF POSSIBLE FLYING TRAINING BEHAVIORS THAT MAY OR MAY NOT HAVE BEEN OBSERVED BY YOU FOR A PARTICULAR STUDENT.*

Assign each pilot candidate a numerical value that reflects his/her performance by writing the value in the space provided below the scale. Scale values range from (1) very poor performance to (9) excellent performance.

After you have assigned a numerical rating, provide at least one behavioral example of pilot candidate flying training activities that supports your rating (e.g., student is constantly unprepared to fly; student is a "team player" - helps other students who are having problems).

1. *Achievement Motivation* - a willingness and determination to work toward goals

| Numerical Rating | | Behavioral Examples |
|---|---|---|
| 9 | | Student is fully prepared for each mission (e.g., reviews readings required for performance of maneuvers for that flight; focuses on the task or assignment; prepared to ask questions about material that is not clear) |
| 8 | More than usual amount of typical activity or effort related to this factor | Student shows willingness to make personal sacrifices to achieve goals (e.g., extra study time) |
| 7 | | Student tries to learn from his/her mistakes (e.g., wants to repeat a maneuver when he/she fails on first attempt) |
| 6 | | Student has well-developed mission profile plan; reviews checklist |
| 5 | Usual amount of typical activity or effort related to this factor | |
| 4 | | Student is attentive during mission debriefings (e.g., takes notes) |
| 3 | | Student is unexcited about training; does not seem to care about his/her performance student is apathetic during mission debriefings, shows up late for debriefings or flights) |
| 2 | Less than usual amount of typical activity or effort related to this factor | |
| 1 | | Student is constantly unprepared for sorties |

Specific examples of this pilot candidates T-37 training activities related to this factor:

_____Numerical rating for this factor

2. *Assertiveness* - Self-assurance (i.e., belief in one's own capabilities) and willingness to defend one's own opinions

| Numerical Rating | | Behavioral Examples |
|---|---|---|
| 9 | More than usual amount of typical activity or effort related to this factor | Student confidently speaks up at Emergency Procedures meetings, presenting his/her opinions in an organized manner |
| 8 | | |
| 7 | | Student asks questions/volunteers to answer Instructor Pilot's questions |
| 6 | Usual amount of typical activity or effort related to this factor | Student is afraid to talk on the radio |
| 5 | | Student challenges IP techniques based on book knowledge or another IP's instructions |
| 4 | | |
| 3 | Less than usual amount of typical activity or effort related to this factor | Student is afraid to talk on the radio |
| 2 | | Student is quiet; does not speak out when classmate says something that is obviously wrong |
| 1 | | Student is easily swayed; does not believe in his/her own ability |

Specific examples of this pilot candidates T-37 training activities related to this factor:

_____Numerical rating for this factor

3. *Cooperativeness* - the willingness to coordinate one's own activities with other people or groups (e.g., crew members)

| Numerical Rating | | Behavioral Examples |
|---|---|---|
| 9 | | |
| 8 | More than usual amount of typical activity or effort related to this factor | Student is a "team player" (e.g., student is willing to exchange duties/take off times with a classmate if this will help the classmate or class as a whole) |
| 7 | | |
| 6 | | |
| 5 | Usual amount of typical activity or effort related to this factor | Student participates in group study sessions |
| 4 | | Student actively participates in class activities (on flight line and socially) |
| 3 | | Student does not participate in class activities (on flight line and socially) |
| 2 | Less than usual amount of typical activity or effort related to this factor | Student does not help other students who are having problems with material; Student is a loner; disinterested in group |
| 1 | | Student isolates himself/herself from others (e.g., studies alone,sits off in a corner away from others) |

Specific examples of this pilot candidate's T-37 training activities related to this factor:

_____Numerical rating for this factor

4. *Stress Tolerance* - the ability to cope with stressful circumstances (in both the short and long term) without significant degradation in performance

| Numerical Rating | | Behavioral Examples |
|---|---|---|
| 9 | | |
| 8 | More than usual amount of typical activity or effort related to this factor | Student is able to maintain simulator and/or aircraft control when faced with unusual or unexpected situations |
| 7 | | |
| 6 | | |
| 5 | Usual amount of typical activity or effort related to this factor | Student is generally calm; difficult to fluster |
| 4 | | |
| 3 | | Student is excited easily |
| 2 | Less than usual amount of typical activity or effort related to this factor | Student becomes confused or "task-saturated" when faced with a large number of tasks |
| 1 | | Student "freezes" at the controls during time critical situations |

Specific examples of this pilot candidate's T-37 training activities related to this factor:

_____Numerical rating for this factor