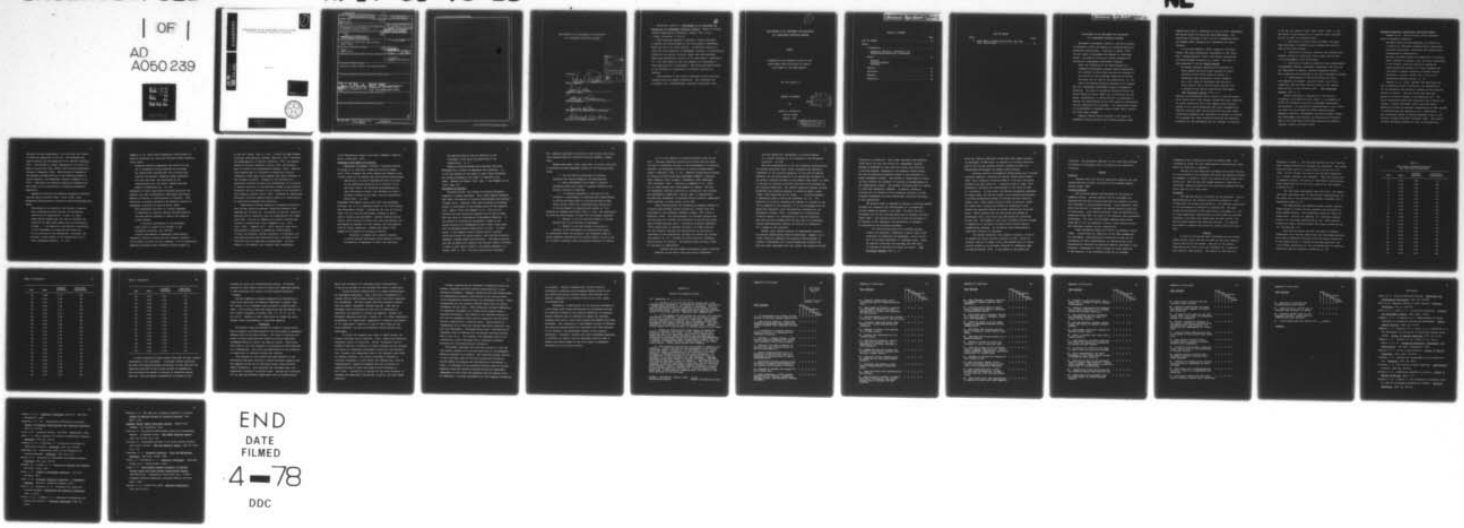
AD-A050 239 AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OHIO F/G 5/9
DEVELOPMENT OF AN INSTRUMENT FOR EVALUATION OF A MANAGEMENT EDU--ETC(U)
AUG 77 R D BALLENTINE

UNCLASSIFIED AFIT-CI-78-21 NL

| OF |
AD
A050 239



END
DATE
FILMED
4-78
DDC

(1)

DEVELOPMENT OF AN INSTRUMENT FOR EVALUATION OF A MANAGEMENT EDUCATION PROGRAM

Ballentine

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: (6) CI 78-21
2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): Development of an Instrument for Evaluation of a Management Education Program.
5. TYPE OF REPORT & PERIOD COVERED: (9) Master's thesis, Thesis
6. PERFORMING ORG. REPORT NUMBER: AFIT-CI-78-21
7. AUTHOR(s): (10) [redacted] Rodger D. Ballentine
8. CONTRACT OR GRANT NUMBER(s): (14)
9. PERFORMING ORGANIZATION NAME AND ADDRESS: AFIT Student at North Texas State University, Denton, Texas
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: AFIT/CI, WPAFB OH 45433
12. REPORT DATE: (11) Aug[redacted]77
13. NUMBER OF PAGES: 34 Pages (12) 41p.
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)
15. SECURITY CLASS. (of this report): Unclassified
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report): Approved for Public Release; Distribution Unlimited
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)
18. SUPPLEMENTARY NOTES: APPROVED FOR PUBLIC RELEASE AFR 190-17.
Jerral F. Guess, Captain, USAF
Director of Information, AFIT
MSGT USAF
Deputy Director of Information
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
012200

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

DEVELOPMENT OF AN INSTRUMENT FOR EVALUATION

OF A MANAGEMENT EDUCATION PROGRAM

APPROVED:

Major Professor

Committee Member

Committee Member

Chairman of the Department of Psychology

Dean of the Graduate School

Ballentine, Rodger D., Development of an Instrument for Evaluation of a Management Education Program. Master of Science (General Experimental Psychology), August, 1977, 34 pp., 1 table, references, 31 titles.

This study was designed to develop a rating instrument to measure the effectiveness of the first phase of management education for an Air Force officer. An officer's ability to lead, the first objective of management training, is intrinsically related to the ability to write, speak, and solve problems. These were behaviorally stated in a 60 item survey. Supervisors (N = 174) were asked to rate the frequency of occurrence of these behaviors for a subordinate. The survey was administered on two occasions to supervisors of officers eligible for training.

Item analysis of the results reflected a strong favorable response bias with usable variability. Data indicated the instrument was a unidimensional internally consistent scale.

(1)

DEVELOPMENT OF AN INSTRUMENT FOR EVALUATION OF A MANAGEMENT EDUCATION PROGRAM

THESIS

Presented to the Graduate Council of the
North Texas State University in Partial
Fulfillment of the Requirements

For the Degree of

MASTER OF SCIENCE

By

Rodger D. Ballentine
Denton, Texas

August, 1977

TABLE OF CONTENTS

LIST OF TABLES

DEVELOPMENT OF AN INSTRUMENT FOR EVALUATION OF A MANAGEMENT EDUCATION PROGRAM

Management development connotes a long-term program of managerial growth and implies an increasing degree of managerial effectiveness. According to Mahler (1953), the end result of management development is individual growth. Building on this point, Mahler discusses the evaluation of management development programs:

> Thus, the evaluation of the effectiveness of a management development program requires measurement of an extent to which there has been an increase in the ability of the individual executive to achieve the goals set for them by the organization. (p. 117)

One of the aspects generally accepted as an integral part of a management development program is managerial education. The topic of managerial education within the United States Air Force (USAF) will be the focus of this study. Specifically, a rating form will be developed to measure the effectiveness of the first phase of professional military education for an officer. The educational program to be considered is Squadron Officer School (SOS), Maxwell Air Force Base, Alabama.

Squadron Officer School consists of 440 hours of residence training covering the following general areas:

Communication Skills, Leadership in the Air Force, Management, and United States Air Force and Force Employment. The educational philosophy of SOS is to be a leadership school for company grade (captains and lieutenants) Air Force officers.

In line with Mahler's (1955) concept of individual growth, SOS seeks professional development of the "whole man," through physical, mental, and ethical disciplines, to achieve maximum potential as a leader. This goal is best expressed in the SOS School Mission:

> To prepare selected captains and lieutenants for those command and staff tasks required of junior officers of the United States Air Force; to strengthen those professional values necessary for a full career of dedication and service to their country; and to provide these officers with a foundation for further professional development. (SOS Curriculum Catalog, 1976, p. 1)

Captains and lieutenants from a variety of jobs and assignments attend one of four classes offered each year, amounting to annual training of over 3,040 company grade officers. Because of the varied background and future assignments of officers attending SOS, course objectives have been established which emphasize the importance of transfer of training. It is assumed that these objectives describe the behaviors necessary for job performance and for transfer of training

to the job (cf. Byars & Crane, 1969; Gilmer, 1966, p. 194). The overall course objective is to improve junior officers' ability to lead, by increasing their

1. Abilities to listen, write and speak, and apply these abilities to logically solve problems that confront them as Air Force officers.

2. Abilities to perceive, evaluate, and redefine their personal concept of leadership, and to apply this concept to the achievement of Air Force goals.

3. Abilities to manage by improving their understanding of management concepts, management techniques, and Air Force and Department of Defense resource management systems, and by applying this knowledge to Air Force management problems.

4. Understanding of the United States Air Force with emphasis on airpower, security issues, total force capability, and how these subjects relate to the company grade officer at the wing/base level. (SOS Curriculum Catalog, 1976, p. 1)

Before beginning an explanation of the specific procedures to be applied to this study, it is necessary to review the literature relating to management development and education. To increase the meaningfulness of this review the topic will be divided into three pertinent areas: Managerial Behavior, Performance, and Effectiveness; Management Development and Training; and Evaluation of Training. Many of the ideas which follow were generated by Campbell, Dunnette, Lawler, and Weick (1970).

Managerial Behavior, Performance, and Effectiveness

Campbell et al. (1970) outline a useful heuristic model of managerial behavior:

> The model specifies that managerial job behavior is a function of individual characterisitics (abilities), volitions (motivation), and social and organizational characteristics (opportunity).
>
> Moreover the model implies the importance of rejecting static concepts of persons, jobs, and their interaction. To be fruitful, practices bearing on managerial effectiveness and research on it must recognize the likelihood of changes occurring in people (through experience, training, growth, etc.), and in the relationship between the two. (p. 16)

Campbell et al. (1970) indicate that describing any job in behavioral terms is difficult, but especially so for the managerial job because it changes so much from one setting to another. Each new setting may mean time, person or situation changes in the managerial job. The direct observation method has often been used to define job content. O'Neill and Kubany (1959) found the direct observation method increasingly inappropriate as jobs increase in sophistication and complexity (e.g., supervision). A more successful method of defining managerial jobs is the Critical Incident Technique (Flanagan, 1954). The Critical Incident Technique consists of a set of procedures for

defining critical requirements of a job which are crucial to effective completion of the job. The technique was used to analyze the job behaviors of AF officers (Flanagan, 1951). This method is useful, especially if an effort is made to account for time, person, and situation determined changes in managerial jobs. Other methods of managerial job analysis include analysis of job dimensions based on factor analysis, and dimensional analysis (Campbell et al. 1970). Describing the behaviors required by managers has been shown to be a prerequisite to adequate performance appraisals.

Adequately defining and measuring managerial effectiveness has been an elusive ideal. Guion (1965) views managerial effectiveness as related to meeting organizational goals:

> The success of an executive lies largely in meeting major organization goals through the coordinated efforts of his organization; in part, at least, these efforts depend upon the kind of influence the executive has upon those whose work his own behavior touches . . . the executive's own behavior contributes to the achievement of organizational goals only by its influence on the perceptions, attitudes, and motives of other people in the organization and on their subsequent behavior. (p. 466)

Campbell et al. (1970) define managerial effectiveness in terms of optimizing all available resources toward organizational goals:

> We define effective managerial job behavior as any set of managerial actions believed to be optimal for identifying, assimilating, and utilizing both internal and external resources toward sustaining, over the long term, the functioning of the organizational unit for which a manager has some degree of responsibility. (p. 105)

Regardless of which organization-oriented definition one selects, the task of developing a meaningful and useable measure (criterion) of effectiveness is difficult. This difficulty is greatly reduced by the heuristic position of conceptual criterion set forth by Astin (1964):

> This conceptual criterion is a verbal statement of important or socially relevant outcomes based on the more general purposes or aims of the organizational sponsor.
>
> A criterion performance is any observable event which is judged to be relevant to the conceptual criterion. (pp. 809-810)

Criteria development of managerial effectiveness frequently involves a search for the ultimate criterion. "The ultimate criterion is not a measure; it is an abstraction, embodying everything that ultimately defines success on

on the job" (Guion, 1965, p. 113). A searth for the ultimate criterion overlooks many problems (Dunnette, 1963), including the dimensionality of criteria (Ghiselli, 1956), the dynamic nature of criteria (Ghiselli & Haire, 1960), and development of criteria for understanding (Wallace, 1965). Research often depends upon the selection of substitute criteria (Thorndike, 1949) whose value depends upon their relevance to the ultimate criterion. Judgment of the relevance of a substitute criterion involves the consideration of its ability to measure only what is intrinsically related to the ultimate criterion, and its inclusion of all sources of the variance in the ultimate criterion (Guion, 1965). Selecting an appropriate substitute criterion measure must be based on a classificatien of organization problems and objectives.

Traditional criterion measures of managerial effectiveness include objective measures, subjective measures, and combinations of these two. As pointed out earlier, objective measures such as production data are inappropriate for higher level jobs. The use of judgments as criteria involves many sources of observational error (Campbell et al., 1970; Guion, 1965). Campbell et al. (1970) identify three major requirements of measures of managerial effectiveness: strongly job centered, devised rationally in accordance with long-range planning and objective-setting, and based on job behaviors that are observable and measurable. The best example of the measure that includes these requirements

is the behaviorally based rating scale (Campbell, Dunnette, Arvey, & Hellervik, 1973).

Management Development and Training

Management development involves a continuous process of change as an individual climbs the executive ladder. Guion (1965) proposes that these changes will depend on the initial traits of the manager and the situations encountered:

> Specifically, it may be hypothesized that the traits of the experienced executive are a function of his own early traits, the variables distinguishing his jobs from those of others, the variables of organizational climates he has known, and the interactions among these. (p. 459)

Management development implies much more than management training, but it is an important aspect of the total management development program. Considering the cost of training today, why are training development programs so widely used in industry and government? Because research results (Miner, 1965) continue to assert that training improves the performance of managers on their present jobs and prepares them for future promotions. McGehee and Thayer (1961) summarize the purpose of training as follows:

> Training in industry has a specific purpose. It should provide experiences which develop or modify the behavior of employees in such a way that what

> the employee does at work is effective in the attainment of the goals and objectives of the organization. (p. 3)

Numerous training methods and techniques have been developed for a variety of employees and situations. It is not the purpose of this paper to cover these techniques. For further information the reader is referred to two books: McGehee and Thayer (1961, chap. 7) and Campbell et al. (1970, chap. 10).

Evaluation of Training

As noted earlier, the history of training evaluation research is clearly favorable. Still there remains disagreement about the measure of training effectiveness and training's ultimate value. Ordiorne (1964) sees training as worthless unless it contributes to the economic goals of the organization. Mahler (1953) and Campbell et al. (1970) do not believe that cost-accounting concepts can be used to determine the value or continuanace of management education. If the results of training are evaluated in terms of its stated objective as recommended by Korb (1956), there could well be agreement between these points of view. In other words, if the company's goals involve profit, training should be oriented to making profits, and training evaluation assesses a change toward this objective. "The basic question that may be asked with regard to any training effort is whether it does in fact yield a change in the people exposed to it" (Miner, 1969, p. 315). The only justification necessary

for a research approach to training is that careful and continuous research makes for effective training (McGehee & Thayer, 1961).

McGehee and Thayer (1961) state that a critical evaluation of training provides management answers to the following questions.

1. Are the dollars being spent on training producing the results needed by the organization?

2. What imrpovements can be made in training procedures which will result in greater returns on the dollars invested in training?

3. Is training necessary in this area or this situation to improvement of organizational effectiveness, or should the money spent on training be used in some other activity which will contribute more effectively to the attainment of organizational goals? (p. 257)

According to McGehee and Thayer (1961, p. 258), the investigator of training-effectiveness is faced with two general questions.

1. Whether or not the training results in desired modification in the behavior of employees.

2. Whether or not the outcome of training is relevant to the achievement of organizational goals.

To investigate these two questions, the researcher needs to secure measures representing the training effect in such a way as to permit inferences about the causal relations of results.

It is to the question of securing measures which we now turn. "The most important problem which arises when one faces the task of evaluating training is the development of a practical, relevant, and measurable criterion of training effectiveness" (Mosel & Tsacnaris, 1959, p. 19). Numerous classification systems of training criteria have been developed (Campbell & Dunnette, 1968; Goodacre, 1955; Lindborn & Osterberg, 1954; MacKinney, 1957; Thorndike, 1949). According to Campbell et al. (1970), all of these systems have three apsects in common: the type of information obtained (objective or subjective); who provides the information (supervisor, trainee, peer, or subordinate); and when the information is obtained (during training, immediately after training, or after return to the job).

Other considerations for developing an adequate measure of training effectiveness include reliability, freedom from bias, practicality, and relevance (McGehee & Thayer, 1961; Tiffin & McCormick, 1965). The characteristic of relevance implies that the training will transfer from the school to the job and attainment of organizational goals. Concerning relevance, MacKinney (1957) stated that "a relevant criterion is a good criterion, one that actually reflects the contributions of the group or individual to the organization and does not contain any extraneous factors" (p. 75). Validity is often cited as a necessary characteristic of criteria. The position taken by Astin (1964) on this point is appropriate:

> Perhaps the most common misconception about criterion measures is the notion that they can be "validated." . . .

> the only method for "validating" a criterion measure is a logical analysis of its relevance to the conceptual criterion. (p. 811)

With due considerations to all the necessary characteristics of criteria mentioned, Astin (1964) pinpointed the essential ingredient of the criterion measures, "In any area of applied research, the criterion measure is an operational statement of the goals or desired outcomes of the program under study" (p. 808). In the area of training evaluation the goals of the program are the training objectives. The logical connection between training objectives and criterion measures is stated by Byars & Crane (1969), "The objectives of the course . . . provide guideposts in developing these criteria and standards" (p. 39). The question of criterion relevance and transfer of training on the job is answered by this approach, if it can be assumed that the objectives adequately cover the behaviors necessary for job performance (Gilmer, 1966). Development of a satisfactory criterion measure is a prerequisite to training evaluation. Once this has been accomplished the next step is to determine through experimentation whether the training is responsible for a change in the criterion.

Several other factors relevant to experimental research on training effectiveness should be mentioned. Campbell et al. (1970) succinctly emphasize the training "effectiveness" is neither a dichotomous nor a single-dimensional variable and that one ideal experiment will not "prove" the program is either

effective or ineffective. Korb (1956) and Mosel and Tsacnaris (1959) point out that the success of a management training program is dependent on more than the course, they held that a training program, regardless of its apparent effectiveness, will not be generally successful unless it also operates in a proper management climate. The success or significance of training also involves the transfer of training to accomplishment of the organization's goals. The question relevance here is a matter or high-level managerial judgment. In general, effective training will be relevant if the conceptual criteria of effectiveness (statements of course objectives) align with the goals of the organization.

The present study is designed to develop a criterion measure suitable for evaluating the effectiveness of a professional military education program in the USAF. The basis for this instrument will be the training objectives of Squadron Officers School (SOS). The training objectives of SOS are developed to be relevant to the subsequent job performance of graudates as indicated by the following statement:

> All young officers must solve problems systematically and logically, communicate clearly, apply sound techniques of leadership and management, and be articulate in the force employment of aerospace power. These are specific abilities and knowledge that SOS seeks to increase in officers who attend the school. (SOS Curriculum Catalog, 1976, p. 3)

Since the training objectives of SOS have been judged relevant to attainment of USAF goals, an analysis of training effectiveness on the job may provide valuable information about the professional development of company grade officers.

Existing criteria of managerial effectiveness were reviewed. Available measures, including the Officer Effectiveness Report (OER), and a composite criterion based on a multiple regression analysis have been ruled out. A previous study by Tupes (1963) attempted to evaluate the effectiveness of SOS using the OER as a criterion. In that study, the OER was considered inadequate. In recent years, the OER has become a non-discriminating measure of officer perfomrance beacuse of inflated ratings (Sturiale, 1968a); and a recent factor analysis study by Sturiale (1968b) has shown that a rating bias related to theofficer's grade affects the appraisal beyond the actual observations of performance. These problems with the OER system have become so predominent that a new Q-sort type OER system was instituted during 1974. Development of any other ultimate criterion through a composition of available variables would involve similar contamination problems. To circumvent these shortcomings a substitute criterion is necessary.

The overall objective, the ultimate criterion, of SOS is to improve the junior officer's ability to lead. Since the officer's ability to speak, write, and problem solve is intrinsically related to the ultimate criterion of leadership (SOS Curriculum Catalog, 1976), it was chosen as the substitute

criterion. The procedures employed in this study were designed to develop an instrument useful in measuring this substitute criterion.

Method

Subjects

Subjects were 174 officers supervising captains and lieutenants who were eligible to attend or had attended Squadron Officer School (SOS).

Criterion Measure

The criterion measure was developed by the method of summated ratings. Content material for the instrument was obtained from the SOS training objectives for speaking, writing, and individual or group problem solving. Sixty behavioral statements were written to represent the desired training objectives in these areas. Since these curriculum areas are differentially emphasized within SOS, the items in the inventory were represented accordingly. The ratio and instrument emphasis among these areas was 3:3:4:1 for speaking, writing, group problem solving, and individual problem solving, respectively.

These statements formed the "Officer Job Behavior Survey" (OJBS). Instructions for completion of the OJBS accompanied the instrument. Supervisors were instructed to rate the job performance of their subordinate(s) by indicating how often subordinate(s) exhibited the positive behavior described in each statement, independent of their ratings for other behaviors. If the behavior in any statements could not be assessed,

respondents were instructed to leave the answer blank. The introductory letter for each administration and behavioral statements are contained in Appendix A.

Ratings for each behavioral statement were scored from one to six according to the frequency with which the behavior occurred. For example, a rating of "never" was scored one and "always" was socred six. When ratings were summed across all items for an individual, behaviors which were not rated were assigned the mean socre value for all items rated.

Procedure

The OJBS was mailed to subjects on two occasions. Initial administration of the inventory occurred on January 27, 1976. On this date 154 surveys were mailed to 116 different supervisors of 154 lieutenants and captains at Lackland AFB. The same survey was distributed to 160 supervisors on September 13, 1976. Of these supervisors, approximately one-third were not surveyed originally, one-third were supervising the same officer, and one-third were supervising different officers. All subjects were asked to complete and return the survey within five workdays.

Results

A total of 314 surveys were distributed and the overall usable survey return rate was 72%--80% for the first administration and 63% for the second. Ratings for 160 surveys, representing the first use of the instrument by a supervisor, were used for item analysis. The results of each item are

displayed in Table 1. The high mean response for each item and small standard deviation of response are noteworthy. The average score was above four on a rating scale of six for 56 of 60 items. Of all ratings, 51% were in the top three categories (i.e. "often, "usually", and "always"). There was variability in summated ratings. The sum of an individual's rating across items ranged from 185 to 360 with a mean of 267 and standard deviation of 32.

The item total correlations were very high. The lowest item total correlation was .32 and the men correlation was .40. The result was a very internally consistent instrument with a coefficient alpha of .97.

The relationship was examined between content subscale items related to either writing, speaking, and individual or group problem solving. The coefficient alphas for these subscales were: writing. 91, speaking .93, individual problem solving .97, and group problem solving .98. These subscale scores were highly intercorrelated--the lowest correlation was .89, the mean was .94.

Data from 58 raters (who were the same or different supervisors of an officer rated during the first administration) were sued to determine interrater and test-retest reliability. On the second survey, 31 ratings by the same supervisor had a test-retest realibility of .89, and scores from 27 different supervisors had an interrater reliability of .84.

Table 1

Item, Means, Standard Deviations, and Item Total Correlations

Item	Mean	Standard Deviation	Item Total Correlations
1	4.81	1.20	.35
2	4.27	1.21	.38
3	4.81	1.04	.36
4	4.12	1.40	.42
5	4.25	1.30	.35
6	4.62	1.20	.36
7	4.60	1.14	.43
8	4.46	1.12	.41
9	3.87	1.46	.45
10	4.57	1.14	.45
11	3.90	1.39	.41
12	4.47	1.27	.33
13	4.68	1.17	.35
14	4.43	1.24	.42
15	4.06	1.46	.50
16	4.62	1.32	.37
17	4.21	1.26	.41
18	4.26	1.24	.41
19	3.97	1.34	.44
20	4.60	1.23	.38
21	4.65	1.30	.40

Item	Mean	Standard Deviation	Item Total Correlations
22	4.48	1.17	.44
23	4.14	1.28	.46
24	4.61	1.23	.41
25	4.67	1.10	.35
26	4.54	1.17	.43
27	4.46	1.23	.41
28	4.51	1.90	.41
29	4.50	1.17	.38
30	4.00	1.40	.41
31	4.37	1.31	.38
32	4.15	1.28	.43
33	4.32	1.36	.32
34	4.07	1.43	.37
35	4.37	1.24	.38
36	4.77	1.07	.37
37	4.48	1.90	.35
38	4.89	1.17	.44
39	4.43	1.16	.39
40	4.88	1.13	.34
41	4.40	1.21	.45
42	4.74	1.18	.39
43	4.49	1.29	.39
44	4.27	1.39	.40

Item	Mean	Standard Deviation	Item Total Correlations
45	4.40	1.31	.39
46	4.36	1.23	.45
47	4.47	1.12	.40
48	4.65	1.24	.41
49	4.53	1.22	.43
50	4.29	1.17	.46
51	4.54	1.33	.43
52	4.18	1.82	.44
53	4.41	1.21	.34
54	4.52	1.19	.46
55	4.43	1.34	.38
56	4.38	1.20	.41
57	4.55	1.16	.42
58	4.46	1.36	.42
59	4.66	1.22	.45
60	4.06	1.39	.45

A factor analysis of these results confirmed the high internal consistency of the instrument. A principal factors extraction was used with squared multiple correlation of each item with the remaining item used as the initial estimate of commonality. Only one factor met Kaiser's criterion of eigenvalue greater than one. This one factor accounted for 93 percent of the

variability within the intercorrelation matrix. No further analysis of these factor analytic results was undertaken because it would duplicate the results of item analysis previously discussed.

Written comments by subjects suggested the elimination of items which duplicated the behavior described in another item. The content of all statements were reviewed to identify duplicates. The following items were repetitious and the behavior was more clearly expressed in another item: 15, 19, 23, 38, 40, 58, and 59. Removal of these items would not change the ratio between items corresponding to curriculum emphasis.

Discussion

The present study was designed to develop a rating instrument to measure the effectiveness of the first phase of management education for an Air Force officer. The intent of training outlined in course objectives was stated in positive behavioral statements forming an officer job behavior survey. Supervisors were asked to rate the frequency of occurrence of these behaviors for a subordinate. The survey was administered on two occasions to supervisors of officers eligible for training.

Item analysis of the results from administration of the performance appraisal reflected a strong favorable response bias. The summated rating for an individual across items reflected usable variability. Data indicated the instrument was a unidimensional internally consistent scale. Ratings were consistent for the same and different supervisors over an 8-month period.

While high reliablity is a desirable scale characteristic, it may be an artifact of the response bias noted in these data.

Strong favorable response bias is a common characteristic of performance appraisals. Such a bias was so predominant in routine officer effectiveness reports that a new Q-sort appraisal was implemented. The data clearly show how favorable response bias can affect "anonymous" research ratings. Even when an anonymous rating eliminated any job risk to subordinate or supervisor the leniency in ratings was apparent. Whether the favorable reaponses noted were a result of: (a) actual subordinate perfomrance; (b) supervisors' expectation of superior performance; or (c) supervisors' inability to rate at scale steps for the traits measured, can only be determined through further scale development.

To reduce the response bias noted in these ratings, the present instrument may be modified. First, repetitious behavioral statements should be eliminated. Second, statements could be worded as positive or negative to reduce response set and increase the chance of independent performance ratings for each behavior. Also, favorable and unfavorable ends of the response scale could be randomly reversed. The present instrument, modified to eliminate redundancy in item content and response set, could be readministered. Remaining response bias can be reduced for summated ratings by using the binary scoring technique of Bass (1956). Variability in ratings for the traits measured is necessary for meaningful reliability, validity, and scale dimensionality.

Further research may be conducted to completely revise the present instrument and force greater discrimination of subordinate performance. Response variability has been demonstrated for behaviorally anchored rating scales or by training raters on the psychometric characteristics of performance ratings. These alternatives are not practical beacuse supervisors are spread throughout the Air Force. Other alternative include redesigning the appraisal instrument as a forced choice between equally favorable attributes. This method is not desirable because of the loss of information about ratee's capability and the difficulties connected with a ipsative scale. Forced ranking of subordinates is not practical because supervisors generally rate only one subordinate. Techniques which force a distribution of responses could be combined with the desired training outcomes used in this study. For instance, the responses could be distributed by a Q-sort technique over a strongest-to-weakest continuum for curriculum-related attributes.

The present instrument has high face validity in that the item content reflects the training objectives of the management education program to be studied. The criterion-related validity of this instrument must be verified by empirical validation studies before the meaning and relationship of the measure is clear. To determine if the logically derived traits of writing, speaking, group and invidiaul problem solving are measurable independent of each other and separable from the general trait of leadership, a revised instrument with item response variability

is necessary. Without response bias, ratings could be correlated to existing job performance measures such as the new forced distribution officer report, peer ratings, and experts' judgments of an officer's ability to write, speak, and problem solve.

Ultimately, a modification of the existing instrument or a new format which forces distribution of performance ratings for behaviors related to the training objecitves is necessary. Further research must be conducted to reduce the favorable response bias noted in this study through the methods previously discussed. Once variability in performance ratings is obtained, the resulting ratings should be correlated with other measures of the officers' job performance related to the training objectives. Although a reliable and valid performance appraisal is difficult to obtain, such an instrument could be used to measure the effectiveness of the first phase of management education for an Air Force officer.

Appendix A

Officer Job Behavior Survey

TO: Supervisor of _____

1. For most of you this will be the second time I have requested your assistance. To determine statistical reliability, a second administration of this survey is necessary. The purpose of and instructions for completion of the attached survey are detailed below. Please detach and return only the completed survey in five work days to: OMYO/Stop 100, Lackland AFB.

2. In accordance with the Privacy Act of 1974, the following information is provided. The authority for collecting information in this survey is contained in 10 USC 8012. The purpose of this inventory is to evaluate the relationship between job performance and Squadron Officer School attendance. Your participation in this survey will provide valuable data. However, if you choose not to participate, no adverse action will be taken.

3. There are certain points which should be emphasized before you complete the survey. First, completion of this inventory will in no way affect the ratee's or supervisor's career and both will remain anonymous; analysis will not be conducted on individual ratings and surveys will be safeguarded to prevent unauthorized access or disclosure. Second, there is a tendency to overuse the above average category on any supervisory rating. Please guard against this tendency.

4. The following instructions are provided to complete this inventory. Please indicate the job performance of the officer identified above in relation to the statements in this survey. Circle the number corresponding to your evaluation for each statement, independent of ratings for other statements. If the behavior in any statement is not applicable or cannot be assessed, do not mark an answer for that statement. Response categories range from "Never" to "Always" to indicate how often this officer exhibited a specific behaivor. When evaluating the frequency of behavior described in each item, use the following reference points for the extreme response categories. The response category "Always" indicates that this officer performed as well as the three best officers of a similar grade you have supervised. Please make any comments about this survey in the space provided on the last page.

RODGER D. BALLENTINE, Captain, USAF
Occupational Analyst

1 Atch
Officer Job Behavior Survey

For office use only
1-8

Answers (9-68)

THIS OFFICER:

	Never	Rarely	Occasionally	Often	Usually	Always
1. is enthusiastic and expends his/her efforts to develop his/her capabilities.	1	2	3	4	5	6
2. when writing formally, constructs supporting paragraphs which begin with a topic sentence that expresses one main idea.	1	2	3	4	5	6
3. is receptive to another person's ideas and tactful in expressing differences of opinion.	1	2	3	4	5	6
4. presents a formal speech, so that his/her physical demeanor adds rather than distracts from his/her purpose.	1	2	3	4	5	6
5. complies with USAF standards for writing style and format outline in AFP 13-2 and AFM 10-1.	1	2	3	4	5	6
6. exerts himself/herself fully in difficult situations and takes full advantage of opportunities to improve himself/herself.	1	2	3	4	5	6
7. approaches problems systematically and gains a thorough understanding of the problem before making a decision.	1	2	3	4	5	6
8. analyzes in advance the purpose of what he/she says.	1	2	3	4	5	6
9. when concluding a formal speech, remotivates the audience, summarizes the main points, and provides a closing statement.	1	2	3	4	5	6

THIS OFFICER:	Never	Rarely	Occasionally	Often	Usually	Always
10. verbally communicates sound conclusions and the reasoning behind them.	1	2	3	4	5	6
12. when speaking formally, captures the audience's attention by motivating them to listen, and highlights the main points.	1	2	3	4	5	6
13. develops means to deal with various situations, and methods to achieve goals.	1	2	3	4	5	6
14. initiates ideas and action when faced with a group problem-solving situation.	1	2	3	4	5	6
15. appears to prepare and rehearse for a formal speech.	1	2	3	4	5	6
16. accepts the responsibility for leading others.	1	2	3	4	5	6
17. when writing formally, clearly and concisely supports the idea expressed in the topic sentence of each paragraph.	1	2	3	4	5	6
18. summarizes the main points and provides a concluding statement when writing a formal letter.	1	2	3	4	5	6
19. organizes written communication by gathering and outlining his/her material.	1	2	3	4	5	6
20. organizes and directs a group toward accomplishment of specific goals.	1	2	3	4	5	6
21. makes decisions with determination and courage.	1	2	3	4	5	6
22. when solving a problem, considers sufficient alternative, and thinks through each to achieve the best solution.	1	2	3	4	5	6

THIS OFFICER:	Never	Rarely	Occasionally	Often	Usually	Always
23. uses relevant, accurate, specific, and adequate support material when speaking.	1	2	3	4	5	6
24. interacts with others by developing a cooperative spirit, resulting in constructive action.	1	2	3	4	5	6
25. when faced with a problem, arrives at a solution that is suitable, flexible, and acceptable.	1	2	3	4	5	6
26. leaves no doubt as to the ideas he/she is expressing through written communication.	1	2	3	4	5	6
27. skillfully and actively participates in helping a group achieve their goals.	1	2	3	4	5	6
28. organizes and writes clearly for a specific purpose.	1	2	3	4	5	6
29. selects a course of action and tranlates it into a practical plan.	1	2	3	4	5	6
30. uses one or more of the following types of support when writing: definitions, examples, comparisons, statistics, quotations, or illustrations.	1	2	3	4	5	6
31. analyzes in advance the purpose of what he/she writes.	1	2	3	4	5	6
32. when writing a formal letter, states the purpose and gives an overview in the introductory paragraph.	1	2	3	4	5	6
33. when confronted with a group problem solving situation, is open to new ideas and free from preconceptions.	1	2	3	4	5	6
34. uses visual aids, when appropriate, to emphasis his/her point when speaking.	1	2	3	4	5	6

THIS OFFICER:	Never | Rarely | Occasionally | Often | Usually | Always					
35. changes so that behaviors, attitudes, etc., conform to new or changed circumstances.	1	2	3	4	5	6
36. verbally responds during conversation in such a way that it is evident he/she has listened carefully.	1	2	3	4	5	6
37. when problem solving, determines the relevance of data in order to select the key factors.	1	2	3	4	5	6
38. has the interest, purpose, enthusiasm, and desire to improve himself/herself.	1	2	3	4	5	6
39. uses simple and direct sentence to express his/her ideas.	1	2	3	4	5	6
40. works in harmony with others for a common purpose.	1	2	3	4	5	6
41. when speaking, provides sufficient main ideas and supporting information to achieve his/her purpose.	1	2	3	4	5	6
42. is willing to listen to the ideas of others when solving group problems.	1	2	3	4	5	6
43. has a comprehensive and exact knowledge of the subject matter he/she is presenting when speaking.	1	2	3	4	5	6
44. uses one or more of the following types of support when speaking: definitions, examples, comparisons, statistics, quotations, or illustrations.	1	2	3	4	5	6
45. communicates sound conclusions and the reasoning behind them when writing.	1	2	3	4	5	6
46. practically and thoroughly uses available means to achieve desired results.	1	2	3	4	5	6

THIS OFFICER:	Never	Rarely	Occasionally	Often	Usually	Always
47. uses correct sentence form and continuity when writing.	1	2	3	4	5	6
48. organizes subordinates and peers and obtains their respect in his/her duty situation.	1	2	3	4	5	6
49. is aware of the needs of the individual within the group, and adequately handles his viewpoints.	1	2	3	4	5	6
50. solves a problem by breaking it into parts, identifying relevant facts and determining their meaning, and projecting the conseuqences of a decision.	1	2	3	4	5	6
51. actively seeks responsibility and is motivated to take action when necessary.	1	2	3	4	5	6
52. when writing a formal letter, includes smooth transitions between sentences and paragraphs.	1	2	3	4	5	6
53. correctly uses numerals, abbreviations, capitalization, and punctuation in his/her writing.	1	2	3	4	5	6
54. applies logical thinking and a systematic method of solving staff-level problems.	1	2	3	4	5	6
55. flexibly and imaginatively devises new solutions or techniques to solve problems.	1	2	3	4	5	6
56. when faced with a problem-solving situation, acts in an independent and fresh way.	1	2	3	4	5	6
57. continually works with and influences a group to achieve their objectives	1	2	3	4	5	6

THIS OFFICER:

	Never	Rarely	Occasionally	Often	Usually	Always
	|	|	|	|	|	|
58. captures my attention and interest when he/she speaks	1	2	3	4	5	6
59. takes action on his/her own as required when working in a group.	1	2	3	4	5	6
60. projects his/her ideas to the audience through a lively sense of communication when speaking.	1	2	3	4	5	6

61. I have supervised this officer for _____months.

COMMENTS:

References

Astin, A. W. Criterion-centered research. Educational and Psychological Measurement, 1964, 24, 807-822.

Bass, B. M. Reducing leniency in merit ratings. Personnel Psychology, 1956, 9, 359-369.

Byars, R. J., & Crane, D. P. Training by objectives. Training and Development Journal, 1969, 23(6), 38-48.

Campbell, J. P., & Dunnette, M. D. Effectiveness of T-group experience in managerial training and development. Psychological Bulletin, 1968, 70, 73-104.

Campbell, J. P., Dunnette, M. D., Arvey, R. D., & Hellervik, L. V. The development and evaluation of behaviorally based rating scales. Journal of Applied Psychology, 1973, 57, 15-22.

Campbell, J. P., Dunnette, M. D., Lawler, E. E., III, & Weick, K. E., Jr. Managerial Behavioral, Performance, and Effectiveness. New York: McGraw-Hill, 1970.

Dunnette, M. D. A note on the criterion. Journal of Applied Psychology, 1963, 47(4), 251-254.

Flanagan, J. C. Defining the requirements of the executive's job. Personnel, 1951, 28, 28-35.

Flanagan, J. C. The critical incident technique. Psychological Bulletin, 1954, 51, 327-358.

Ghiselli, E. E. Dimensional problems of criteria. Journal of Applied Psychology, 40(1), 1-4.

Ghiselli, E. E., & Haire, M. The validation of selection tests in light of the dynamic character of criteria. Personnel Psychology, 1960, 13, 255-231.

Gilmer, B. v. H. Industrial Psychology (2nd ed.). New York: McGraw-Hill, 1966.

Goodacre, D. M., III. Experimental evaluation of training Journal of Personnel Administration and Industrial Relations, 1955, 2, 143-149.

Guion, R. M. Personnel Testing. New York: McGraw-Hill, 1965.

Korb, L. D. How to measrue the results of supervisory training. Personnel, 1956, 32, 278-291.

Lindborn, T. R., & Osterberg, W. Evaluating the results of supervisory training. Personnel, 1954, 31, 224-228.

MacKinney, A.C. Progressive levels in the evaluation of training programs. Personnel, 1957, 34, 72-77.

Mahler, W. R. Evaluation of management development programs. Personnel, 1953, 30, 116-122.

McGehee, W., & Thayer, P. W. Training in Business and Industry. New York: Wiley, 1961.

Miner, J. B. Studies in Management Education. New York: Springer, 1965.

Miner, J. B. Personnel Industrial Relations: A Managerial Approach. New York: MacMillan Company, 1969.

Mosel, J. N., Tsacnaris, H. J. Evaluating the supervisor training program. Engineering and Industrial Psychology, 1959, 1, 18-23.

O'Neill, H. E., & Kubany, A. J. Observation methodology and supervisory behavior. Personnel Psychology, 1959, 12, 85-96.

Odiorne, G. S. The need for an economic approahc to training. Journal of American Society of Training Directors, 1964, 18(3), 3-12.

Squadron Officer School Curriculum Catalog. Maxwell AFB, Alabama: Air University, 1976.

Sturiale, G. The Officer Affectiveness report as a performance mseaure: A research review. USAF AFHRL Technical Report, 1968, No. 68-113, 21 p. (a).

Sturiale, G. Stereotyped patterns in Air Force officer performance factor ratings. USAF PRL Technical Report, 1968, No. 68-6, 22 p. (b).

Thorndike, R. L. Personnel Selection: Tests and Measurement Technique. New York: Wiley, 1949.

Tiffin, J., & McCormick, E. J. Industrial Psychology. Englewood Cliffs, N.J.: Prentice-Hall, 1965.

Tupes, E. C. Relationship between attendance at Squadron officer school and later officer effectiveness reports (PRL-TDR-63-10). Lackland Air Force Base, Tex.: 6570th Personnel Research Laboratory, Aerospace Medical Division, April, 1963.

Wallace, S. R. Criteria for what? American Psychologist, 1965, 20, 411-417.