DOCUMENT RESUME

ED 395 965 TM 025 103

AUTHOR Stricker, Lawrence J.
TITLE Assessing Leadership Potential at the Naval Academy with a Biographical Measure.
INSTITUTION Educational Testing Service, Princeton, N.J.
SPONS AGENCY American Society for Engineering Education, Washington, D.C.; Army Research Office, Durham, N.C.; Department of the Navy, Washington, D.C.; Navy Personnel Research and Development Center, San Diego, Calif.
REPORT NO ETS-RR-89-14
PUB DATE Mar 89
NOTE 67p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Achievement Need; *Biographical Inventories; *Educational Assessment; Emotional Development; Higher Education; Interpersonal Competence; Item Analysis; *Leadership; Measurement Techniques; *Military Personnel; Personality Traits; *Test Construction; Test Validity
IDENTIFIERS *Naval Academy MD; Navy

ABSTRACT

The aim of this study was to construct and validate a biographical inventory to measure personality traits that are predictive of leadership. The experimental inventory, consisting of tentative scales for Dominance, Emotional Stability, Need for Achievement, Self-Confidence, and Sociability, was administered to 642 incoming midshipmen at the Naval Academy. The tentative scales were item analyzed, and the validity of the final scales was appraised against subsequent peer ratings of the traits and of leadership, as well as against first-semester grades that may reflect leadership (Military Performance and Professional Military Quality Point Rating). The Sociability, Dominance, and Need for Achievement scales demonstrated convergent and discriminant validity and correlated with the leadership criteria; the Emotional Stability and Self-Confidence scales showed no signs of validity. Although the level of validity was often modest, the Sociability scale correlated appreciably (0.28) with the peer rating of leadership, suggesting that this measure may have practical value in assessing leadership potential. (Contains 1 figure, 15 tables, and 16 references.) (Author/SLD)

RESEARCH REPORT

ASSESSING LEADERSHIP POTENTIAL AT THE NAVAL ACADEMY WITH A BIOGRAPHICAL MEASURE

Lawrence J. Stricker

Educational Testing Service
Princeton, New Jersey
March 1989

Assessing Leadership Potential at the Naval Academy with a Biographical Measure

Lawrence J. Stricker

ETS Research Report

March 1989

Abstract

The aim of this study was to construct and validate a biographical inventory to measure personality traits that are predictive of leadership. The experimental inventory, consisting of tentative scales for Dominance, Emotional Stability, Need for Achievement, Self-Confidence, and Sociability, was administered to incoming midshipmen at the Naval Academy. The tentative scales were item analyzed, and the validity of the final scales was appraised against subsequent peer ratings of the traits and of leadership, as well as against first-semester grades that may also reflect leadership (Military Performance and Professional Military Quality Point Rating). The Sociability, Dominance, and Need for Achievement scales demonstrated convergent and discriminant validity, and correlated with the leadership criteria; the Emotional Stability and Self-Confidence scales showed no signs of validity. Although the level of validity was often modest, the Sociability scale correlated appreciably (.28) with the peer rating of leadership, suggesting that this measure may have practical value in assessing leadership potential.

Assessing Leader nip Potential at the

Naval Academy with a Biographical Measure

Applicants to the Naval Academy are intensively screened. A variety of sources of information are used for that purpose, including scholastic aptitude tests, an interest inventory, the high school record, recommendations, and an interview. In recent years, the focus of the screening was on identifying applicants who would be successful students. However, in 1984, the Secretary of the Navy directed that the emphasis be shifted to identifying applicants who would be future leaders. This study represents one effort in this direction: the development of a biographical measure to assess the leadership potential of Academy applicants.

A biographical measure was chosen for this purpose because such devices have distinct advantages over personality and interest inventories, and similar instruments. Biographical measures capture directly the past behavior of a person, probably the best predictor of his or her future actions. And the measures deal with facts about the person's life, not the introspections and subjective judgments that make up the content of personality inventories and the like. As a result, biographical measures are likely to be less prone to misinterpretation, resistance, and distortion.

Many current biographical measures have serious limitations, limitations that are not inevitable and that prevent these devices from realizing their full potential (Stricker, 1987). Their items may not be true biographical items: items that deal with factual matters and are answerable in a factual way. Instead, the items may concern opinions or internal states, and answering the items may require

subjective judgments. Using such nonfactual items ignores the unique virtues of biographical measures and dissipates their strengths. The items may also concern behavior by the examinees and other matters not under their control. Such items are problematic, both ethically and psychometrically: they are unfair and invalid, and highly susceptible to sex, ethnic group, and similar kinds of bias. The psychometric procedures used in developing biographical measures are also troublesome. The common practice of using empirical keying results in scores that are difficult to interpret and that may not be applicable to other situations. A related problem is that the sole focus on maximizing convergent validity entirely neglects the discriminant validity of the measures, including the measures' contamination with response styles or with sex or race bias.

These difficulties are not inherent to biographical inventories. General purpose biographical measures can be developed by combining homogeneous sets of items in the same way that personality and interest inventories have long been constructed. An early example is Siegel's (1958) Biographical Inventory for Students. And Jackson (1970), in developing the Personality Research Form, delineated test construction procedures that can maximize both the convergent and discriminant validity of an inventory.

Recent reassessments of the empirical research on the personality correlates of leadership ability suggest that this ability can be successfully predicted (Aronoff & Wilson, 1985; Bass, 1981; Hogan, 1987; Lord, De Vader, & Alliger, 1986). Consistent links appear to exist between personality traits and leadership. This reassessment

calls into question the widely-held conclusion, stemming from influential reviews by Stogdill (1948) and Gibb (1954), that personality traits and other individual-difference variables do not distinguish leaders from followers, and that leadership is simply a function of the situation. This reappraisal of the work on leadership also raises the real possibility that a properly constructed biographical inventory, designed to measure relevant personality traits, may be able to assess leadership potential.

The main purpose of this study was to construct and validate a biographical inventory to measure personality traits that are predictive of leadership. A secondary goal was to demonstrate the feasibility and value of constructing an inventory that is free of the limitations of current biographical measures.

Method

Overview

A biographical inventory made up of five tentative personality scales and a Social Desirability scale was assembled. The inventory was group administered on July 2, 1987, the second day of Plebe Summer, to 642 entering midshipmen at the Academy. These were all the midshipmen in the starboard battalion, except for a small number standing watch or otherwise required to be absent. (Plebe Summer is a seven-week bootcamp for the training and indoctrination of incoming midshipmen.) The inventory was the first of several devices (the others were personality scales and inventories) administered to the midshipmen in two testing sessions. The midshipmen were instructed that some of the questionnaires were being administered for research

purposes and others for individual counseling, and the results would not become part of the midshipmen's official records.

Seven weeks later, on August 14 to 17, 1987, the last weekend of Plebe Summer, the same battalion of midshipmen rated all the members of their squad on leadership and on the same personality traits assessed by the inventory. Data on other variables were secured from admissions records, and first-semester grades were subsequently obtained.

The five foreign nationals who had been administered biographical inventories were excluded from the study, and the remaining midshipmen were divided into three subsamples: one random half (N = 233) of 1987 high school graduates, the other half (N = 233) of 1987 graduates, and earlier graduates (N = 171).

The item analysis of the inventory was carried out for one random half of 1987 graduates, and the analysis of the inventory's validity was done for the two other samples.

Constructing the Inventory

Defining and assembling tentative personality scales. The empirical research on the personality correlates of leadership was reviewed. Because this literature is massive and has been extensively reviewed, the reviews themselves were reviewed (Aronoff & Wilson, 1985; Bass, 1960; Gibb, 1954, 1969; House & Baetz, 1979; Mann, 1959; Stogdill, 1948). Despite differences in terminology used in the reviews and in the reports cited, five personality traits that can be characterized as dominance, emotional stability, need for achievement, self-confidence, and sociability were identified as being more-or-less

consistently related to leadership in the reviews.[1] The results of the present review are summarized in Table 1.

Insert Table 1 about here

Items were written to tap each of the five traits, as they were defined by Dermen, French, and Harman (1978) in their synthesis of the factor analytic literature in the personality domain. The definitions of these traits appear in Table 2. Six people contributed items. Ideas for items were gleaned from existing personality, interest, and biographical measures; existing published and unpublished data on relevant behaviors identified in act frequency research (Buss & Craik, 1981); and new data on relevant behaviors, obtained with act frequency methods at the Academy and four other schools.

Insert Table 2 about here

The items have these characteristics:

1. The items use multiple-choice or Yes-No formats.

2. The alternatives for multiple-choice items are on continuous scales. This feature facilitates quantification of the responses.

3. The items (stems and alternatives) are factual. This characteristic accords with the basic conception of a biographical item.

4. The items deal with public behavior. (Statements that the examinees made to others are included because the statements reflect internal states that are critical aspects of self-confidence and need for achievement.) The factual nature of the items, in combination

with the public behavior inquired about, makes the responses, in principle, verifiable; the verifiability is expected to reduce distortion.

5. The items (a) concern behavior by the examinees that is under their control and (b) involve opportunities and resources available to virtually everyone. These characteristics are expected to enhance the validity of the items, while minimizing unfairness and bias.

6. Apart from a few items dealing with easily recalled activities that inquire about whether these things occurred at any time in the examinees' lives, most items concern a particular time period: during high school, during the senior year, or since entering the senior year. This practice standardizes the period being described, and focuses on a recent period that can be recalled accurately and is likely to be most relevant to the examinees' current behavior.

The items went through a series of pilot tests; reviews for clarity, appropriateness, and relevance by panels of psychologists and by Academy staff; and subsequent revisions.

1. The initial set of 276 items was pilot tested with seven undergraduates or graduate students on the temporary staff at the Navy Personnel Research and Development Center (NPRDC), and reviewed by seven NPRDC psychologists. Items were revised or dropped on the basis of the pilot test and reviews, and new items were added.

2. The new set of 371 items was pilot tested with six college freshmen; and reviewed for clarity, objectionability, and validity by an Educational Testing Service (ETS) psychologist with expertise in

test development and personality. The items were also classified by three judges with expertise in personality and clinical psychology (an ETS psychologist and two advanced graduate students in clinical psychology at the New School for Social Research) with regard to the items' relevance to the traits, using the Dermen et al. trait definitions.

The key instructions to the judges follow:

> Please read each item and decide whether it appears to be primarily a measure of one of the five traits. (Use the attached definitions of the trait, as far as possible, but they may need to be supplemented by your judgment.) If an item appears to be primarily a measure of a particular trait, put it in the pile for that trait. If the item does not appear to be primarily a measure of any of the traits or appears to measure two or more of the traits more-or-less equally well, put it in an "Other" pile....

Based on the judgments, items were grouped as follows:

Substantial agreement (i.e., at least three of the four people, including the item writer, who classified the item put it in the same trait category): Two or three judges classified the item in the intended trait category, or all three judges put the item in the same unintended trait category. In the latter case, the item was then reassigned to the unintended category.

Borderline agreement (i.e., two of the people put the item in the same trait category, and two put the item in two other categories): One judge classified the item in the intended trait category, and the

two other judges classified the item in two different unintended trait categories or the "other" category; or two judges classified the item in the same unintended trait category, and the third classified the item in another unintended trait category or the "other" category. In the latter case, the item was then reassigned to the unintended category.

No agreement (i.e., two or more people did not put the item in the same trait category, or two put the item in one category but two put it in another category): The three judges classified the item in three different unintended trait categories or the "other" category; one judge classified the item in the intended trait category, and the other two classified the same unintended trait category or the "other" category; or all three judges classified the item in the "other" category. In any of these cases, the item was then dropped.

Items were revised or eliminated on the basis of the pilot testing, review, and judging. Only minor revisions were made in order to avoid changing the meaning of the items, as determined by the judging.

3. The surviving set of 257 items, assembled into an inventory, with items arranged in proper order and instructions added, was reviewed by the Academy staff. Items were changed or dropped on the basis of the review. Yes-No items were also revised, as necessary, so that they were roughly balanced in their keying on each scale (35% to 65% of the items keyed Yes) in order to minimize the effects of acquiescence. Revisions were limited to avoid changing the meaning of the items.

4. A revised version of the inventory was pilot tested with four students (a high school senior and three undergraduates). Minor revisions were made in the items, and additional items were dropped.

Assembling a tentative social desirability scale. Forty-six additional items were written, similar in form to the items for the personality scales: two items for each of the 23 personality factors that Dermen et al. identified, beyond the five factors included in the inventory. The items were intended to be clearly desirable or undesirable.

The items were pilot tested, with the items for the personality scales, with the six college freshmen and with the other four students. The items were also reviewed, with the items for the personality scales, by the Academy staff. In addition, the items were administered to 79 undergraduates enrolled in an introductory psychology course in order to identify items on which consensus about the items social desirability exists. Usable data were obtained for 70 students.

The key instructions follow:

> Suppose that the attached questionnaire were being used to decide who should be admitted to one of the armed services academies--the Naval Academy ("Annapolis"), the Military Academy ("West Point"), or the Air Force Academy. Which answers would probably be considered desirable by the admissions office?
>
> ...select the answers that you think the admissions office would consider desirable, not the answers that you

personally consider desirable or the answers that actually describe you....

Yes-No items were dropped if the proportion of socially desirable responses was not significantly different from .50 ($p > .05$, two-tail) Other items were revised or eliminated on the basis of the pilot testing and review. As far as possible, only minor revisions were made to avoid changing the meaning of the items. Thirty-six items survived.

Instructions. The final inventory consisted of 240 items for the tentative personality scales made up of items with clear and borderline agreement and 36 items for a tentative Social Desirability scale. The instructions emphasized the factual nature of the items and their verifiability in an effort to minimize distortion.

The key instructions follow:

> This inventory consists of a number of factual questions about your activities and experiences. Answer the questions as accurately as possible; your answers may be verified for their accuracy.

Item analysis of personality scales. The item analytic procedures were modeled after and extended those described by Jackson (1970). The modified procedures were intended to (a) maximize convergent and discriminant validity; and (b) ensure that response styles, and sex and ethnic group bias, were minimized.

The item analysis was carried out for a random half of the 1987 graduates (N = 233) Eighty-nine percent of the sample was male and the same percent were White.

The number of items analyzed ranged from 42 to 60 for the five personality scales. The item analysis focused on each item as a whole, not on individual response alternatives. Alternatives for multiple-choice items were dichotomized at the median. Items were scored 0 or 1, with all responses in the same dichotomy being assigned the same score.

1. All items with extreme endorsement frequencies (less than .05 or greater than .95) were eliminated.

2. The following correlations were computed for each of the remaining items:

a. Its correlation with the total score for its own scale.

b. Its correlation with the corresponding total score for each of the other personality scales.

c. Its correlation with the total score for the Acquiesence scale (the latter consists of 17 items from the tentative personality scales and is balanced in content and keying; its Coefficient Alpha reliability was .63).[2] This correlation was only computed for Yes-No items in the 169-item section of the inventory made up exclusively of such items.

d. Its correlation with the total score for the Social Desirability scale (the latter consists of 12 items and is balanced in content and keying; its Coefficient Alpha reliability was .46).[3]

e. Its correlation with sex (Male = 1, Female = 0).

f. Its correlation with ethnicity (White = 1, All others = 0).

All correlations were product-moment indexes (i.e., point-biserial correlations or phi coefficients). Correlations of items

with their own scale were corrected for overlap when the item was included in the total score, correlations of Yes-No items with the Acquiescence scale were corrected for overlap when the item was included in that scale, and all correlations were corrected for attenuation in the total scores (using the Coefficient Alpha reliability of the scores).

3. An item was eliminated if:

a. It did not correlate significantly ($p > .05$, one-tail) with its own scale (the significance test was applied to the actual correlations, not the attenuation-corrected correlations; Bobko & Rieck, 1980).

b. It had a correlation with one of the other personality scales, the Acquiescence scale (if a Yes-No item), the Social Desirability scale, sex, or ethnicity that equaled or exceeded the correlation with its own scale.

The item analysis was done twice, and all the items were included in each analysis. In the first analysis, the score for each personality scale was based on the set of items for which there was substantial agreement among the judges. (Seven of the Yes-keyed items on the Need for Achievement scale were randomly excluded from the total score for that scale so that the items in the score were roughly balanced in their keying--no more than 65% of the items are Yes-No items keyed in the same direction.) The number of items in the total scores ranged from 21 to 38, and the other items analyzed for each scale range from 12 to 28. At the end of this analysis, the number of surviving items ranged from 14 to 21 per scale.

In the second analysis, the score for each personality scale was based on all the items that survived the initial analysis. Three of the surviving Yes-keyed items on the Dominance scale and ten of these Yes-keyed items on the Self-Confidence scale, those with the lowest correlations with their total scores in the initial analysis, were excluded from the total score to balance the keying.) The number items in the total scores ranged from 8 to 21, and the other items analyzed ranged from 21 to 52 per scale. The number of items that survived this analysis ranged from 14 to 22 per scale.

The final scales were based on the items that survived the second analysis. (Eight of the surviving Yes-keyed items on the Self-Confidence scale, those with the lowest correlations with their total score in the second analysis, were excluded from the total score in order to achieve balanced keying.) The number of items on the scales ranged from 8 to 22. The number of items at each step are reported in Table 3.

Insert Table 3 about here

Validity Analysis

The validity analysis was intended to assess the ability of the final forms of the personality scales to tap leadership as well as the personality traits that they were intended to measure. (It is conceivable that a scale may validly measure the intended trait but be unrelated to leadership, at least as it is manifested in the setting being studied.) This analysis was also designed to appraise the involvement of response styles in the personality scales, sex and

ethnic group differences on the scales, and the overlap between the scales and current admissions measures.

The analysis was done separately for the random half (N = 233) of 1987 graduates not used in the item analysis and for earlier graduates (N = 171). Of the 1987 graduates, 91% were male and 85% were White; the corresponding percentages for the earlier graduates were 92% and 81%. Of the earlier graduates, 43% had attended the Naval Academy Preparatory School (NAPS), 20% had participated in the U. S. Naval Academy Foundation program, and 32% had attended college. It was anticipated that the inventory might be less valid for earlier graduates because many items on the inventory deal with the period since the midshipmen entered senior year of high school, and this period of time varies in quantity and quality for earlier and recent graduates.

Criteria. Peer ratings were the criteria for the personality traits, and the primary criterion for leadership. Secondary criteria of leadership were two Academy grades, Military Performance and Professional Military Quality Point Rating (MQPR), which may reflect leadership as well as other variables. Two other Academy grade criteria, unrelated to the personality traits and leadership, were included in the analysis for exploratory purposes: Conduct and Academic Quality Point Rating (AQPR).

The criteria follow:

1. Peer ratings for Dominance, Emotional Stability, Need for Achievement, Self-Confidence, Sociability, and Leadership. All the midshipmen in the starboard battalion were asked to rate themselves

and the other members of their squad (most squads had 12 or 13 midshipmen) on the five personality traits plus leadership. The poles of the variables were defined, and an eight-point scale was used. The trait definitions were adapted from Dermen et al., and the leadership definition from Gibb (1954). The rating scales and their scoring are shown in Figure 1.

Insert Figure 1 about here

The key instructions follow:

> The information requested on these forms is for research purposes only. It will be seen only by authorized Department of Navy research personnel who will use the information for statistical analyses. Nobody at the Naval Academy and no one else in the Navy will have access to it. And it will not become part of your own military record or the military record of others. When you have completed the forms, put them in the attached envelope, seal it, and give it only to the officer in charge. Do not put name, alpha [identification] number, or other identifying information on the forms or the envelope.
>
> Your task is to rate each of the members of your squad, including yourself, on several characteristics....You are to record your ratings on these forms, using an 8-point scale....
>
> Make the ratings as accurately as possible, and try to use all of the points on the scale....

The mean number of raters per squad was 11. A total of 654 midshipmen were rated. Because of clerical errors, a few midshipmen were not rated at all and the midshipmen in one squad were not rated for Leadership. Ratings were excluded for midshipmen who were foreign nationals (N = 6) or who left their squads, because of attrition or transfer, before the end of Plebe Summer (N = 25).

The mean number of raters contributing to the ratings of each remaining midshipmen was 11 for each of the six ratings. The median rating received by each midshipman was calculated for the total number of raters in the squad as well as for two random halves of these raters. (Because the ratings were anonymous, self-ratings could not be eliminated; the median was used to reduce the influence of deviant self-ratings.) The medians were standardized within squads and transformed to standard scores (means of 50 and standard deviations of 10).

2. Military Performance--first semester (A = 4...F = 0; a grade assigned by the company officer, based on performance in a variety of areas, including drill and parades, standing watch, sports participation, extracurricular activities, personal appearance and military bearing, room appearance, Conduct grade, and academic performance--including Physical Education.)

3. MQPR--first semester (an average of grades for Military Performance, Conduct, Physical Education, and professional courses).

4. Conduct--first semester (A = 4...F = 0; based on demerits received for violating regulations concerning midshipmen's conduct).

5. AQPR--first semester (an average of grades in academic courses).

Other variables. Other variables in the analysis were sex, ethnicity, and response style scales, as well as measures currently used in admissions. These variables follow:

1. Sex (Male = 1, Female = 0)

2. Ethnicity (White = 1, All others = 0).

3. Acquiescence scale (an 11-item scale, balanced in content and keying; it differs from the Acquiescence scale used in the item analysis of the personality scales).[4]

4. Social Desirability scale (a 12-item scale, balanced in content and keying; it differs from the Social Desirability scale used in the item analysis of the personality scales).[5]

5. Average Scholastic Aptitude Test/American College Testing Program (SAT/ACT) Verbal score.

6. Average SAT/ACT Mathematical score.

7. Rank in high school class (converted percentile).

8. Recommendations of high school officials (adjusted).

9. Extracurricular activities in high school (ECA)-Athletic.

10. ECA-Nonathletic.

11. Combined ECA (a combination of ECA-Athletic and ECA-Nonathletic).

12. Predicted AQPR (a composite of SAT/ACT scores and Rank in High School Class)

13. Candidate Multiple (adjusted; a composite of SAT/ACT scores, Rank in High School Class, Recommendations of High School Officials, Combined ECA, and specially-developed Strong-Campbell Interest Inventory measures: Engineering/Science or Humanities, and Career Retention).

Results and Discussion

Intercorrelations of Personality Scales, Response Style Scales, Ethnicity, and Sex

The intercorrelations of the personality scales, response style scales, sex, and ethnicity for the two samples are reported in Table 4 and 5. The Coefficient Alpha reliabilities of the personality and response style scales are also shown in these tables.

Insert Tables 4 and 5 about here

The results were similar for both samples. The personality scales were moderately reliable (.66 to .78 for 1987 graduates, .62 to .77 for earlier graduates), with Sociability being appreciably more reliable than the other scales (.78, .77). The Acquiescence scale was somewhat less reliable (.58, .57), and the Social Desirability scale was minimally reliable (.27, .35). The moderate reliability of the personality scales and the Acquiescence scale is not surprising in view of their short length. The marginal reliability of the Social Desirability scale clearly limits the value of this measure.

All the scales correlated positively or near zero with each other, except for Self-Confidence, which correlated negatively with Need for Achievement (-.22, -.27). Dominance and Sociability correlated substantially with each other (.57, .51), and Need for

Achievement correlated moderately with Dominance (.39, .36) and Sociability (.44, .28).

The Emotional Stability and Self-Confidence scales correlated moderately with the Acquiescence scale (-.25, -.26 for the former; -34, -.28 for the latter). And the Emotional Stability, Need for Achievement, and Sociability scales correlated moderately with the Social Desirability scale (.28, .35; .36, .35; and .37, .28 for the three scales, respectively). These correlations imply some response style involvement in the personality scales, despite the precautions taken in the item analysis. The correlations for the Social Desirability scale underestimate the involvement of this response style, given the low reliability of the scale.

All the scales correlated near zero with sex and ethnicity, suggesting that bias is absent from these scales.

Intercorrelations of Criteria

The intercorrelations of the criteria are shown in Tables 6 and 7. The interrater reliabilities of the ratings also appear in these tables. (Interrater reliability was estimated by the correlation between the ratings, in standard score form, for the random halves of raters, corrected for double length by the Spearman-Brown formula.)

Insert Tables 6 and 7 about here

The results were similar for both samples, except for the marginally higher correlations of the peer ratings with Military Performance and MQPR for 1987 graduates. All the peer ratings were highly reliable (.76 to .89 for 1987 graduates, .81 to .87 for earlier

graduates). In general, the ratings correlated highly with each other (.44 to .84 for 1987 graduates, .36 to .86 for earlier graduates), though Sociability correlated only moderately with Emotional Stability (.44, .36) and Need for Achievement (.45, .39). The appreciable intercorrelations raise the possibility that a halo factor is inflating them.

All the peer ratings generally correlated moderately with Military Performance (.38 to .52 for 1987 graduates, .23 to .41 for earlier graduates) and MQPR (.37 to .51, .23 to .42), except Sociability, which correlated slightly with these criteria (.26, .17 for Military Performance; .16, .10 for MQPR). Need for Achievement correlated markedly higher with these criteria (.52, .41 for Military Performance; .51, .42 for MQPR) than did the other ratings, including Leadership (.44, .32; .39, .29).

Military Performance and MQPR correlated highly with each other (.85, .89), reflecting the part-whole relationship between these variables.

Need for Achievement and Military Performance correlated moderately with AQPR (.26, .23 for the former; .42, .32 for the latter), and MQPR correlated moderately to substantially with Conduct (.52, .25). The correlations of Military Performance with AQPR, and of MQPR with Conduct, also reflect the common components in the variables.

Intercorrelations of Peer Ratings for Random Halves of Raters

In order to explore the correlations among the peer ratings in

more detail, the intercorrelations of the ratings for two random halves of the raters are given in Tables 8 and 9.

Insert Tables 8 and 9 about here

The results were generally similar for the two samples, but some difference did occur. For 1987 graduates, in the case of Dominance, Emotional Stability, Self-Confidence, and Leadership, the correlations between the ratings of the same construct by different raters (e.g., the two groups' ratings of Self-Confidence correlated .76) were equaled or exceeded by the correlations between ratings of these constructs and ratings of different constructs by the same raters (e.g., one group's ratings of Self-Confidence correlated .81 with its ratings of Dominance). This result--excessive correlations between ratings of different constructs by the same raters--is an indication of method variance in the ratings, presumably a halo factor.

Furthermore, in the case of Emotional Stability and Self-Confidence, the correlations between the ratings of the same construct by different raters (e.g., .76 for Self-Confidence) were equaled or exceeded by correlations between ratings of these constructs and ratings of different constructs by different raters (e.g., one group's ratings of Self-Confidence correlated .78 with the other group's rating of Leadership). This result--excessive correlations between ratings of different constructs by different raters--is another sign of limited discriminant validity of the ratings, unrelated to method variance.

For earlier graduates, in the case of all constructs except Sociability, the correlations between the ratings of the same construct by different raters were equaled or exceeded by the correlations between ratings of these constructs and ratings of different constructs by the same raters. Like the parallel result for 1987 graduates, this is a sign of method variance.

In sum, for 1987 graduates, five of the seven ratings (all except Need for Achievement and Sociability) appear to have limited discriminant validity, primarily because of the presence of method variance. For recent graduates, only one of the ratings (Sociability) does not have limited discriminant validity due to method variance.

Correlations of Personality Scales with Criteria

The correlations of the personality scales with the criteria are reported in Tables 10 and 11.

Insert Tables 10 and 11 about here

In general, the scales had greater validity for 1987 graduates. For this sample, the Dominance and Sociability scales correlated moderately (.23, .36), and the Need for Achievement and Self-Confidence scales correlated slightly (.12, .12), with the corresponding peer rating. However the Self-Confidence scale correlated higher with an irrelevant rating than with the corresponding rating.

The Sociability scale correlated moderately (.28) with the Leadership rating. And the Sociability and Need for Achievement

scales correlated slightly with Military Performance (.13, .12) and MQPR (.12, .13).

The Sociability scale also correlated slightly (.16) with Conduct.

For earlier graduates, the Dominance and Sociability scales correlated moderately with the correspondirg peer rating (.20, .39), but the Dominance scale correlated higher with an irrelevant rating.

The Dominance and Sociability scales correlated slightly with the Leadership rating (.14, .16).

The Sociability scale correlated slightly--and negatively--with AQPR (-.17).

In short, there was some evidence of convergent and discriminant validity for three scales (Dominance, Need for Achievement, and Sociability) for 1987 graduates, and one scale (Sociability) for earlier graduates. Furthermore, one scale (Sociability) had some validity in predicting leadership, at least for 1987 graduates. The greater validity of the personality scales for 1987 graduates was expected because of the nature of the inventory content.

Correlations of Personality Scales with Other Predictors

The correlations of the personality scales with the other variables are given in Tables 12 and 13.

Insert Tables 12 and 13 about here

The correlations were consistently higher for the 1987 graduates. For this sample, the Sociability scale correlated moderately with ECA-Athletic (.31), ECA-Nonathletic (.32), Combined ECA (.47), and Average

SAT/ACT Verbal and Mathematical scores (-.28, -.35)--negatively with the latter scores. The Dominance scale correlated moderately with ECA-Nonathletic (.39) and Combined ECA (.39).

For earlier graduates, the Dominance scale correlated moderately (.29) with Combined ECA. (Data on the correlations for ECA-Athletic and ECA-Nonathletic are not available for this sample.)

In short, the Sociability scale overlapped with ECA and SAT/ACT scores for 1987 graduates--negatively for the latter, and the Dominance scale also overlapped with ECA scores for both samples.

Correlations of Other Predictors with Criteria

The correlations of the other predictors with the criteria appear in Tables 14 and 15.

Insert Tables 14 and 15 about here

The correlations were consistently higher for 1987 graduates. For this sample, several variables correlated moderately with the leadership criteria: ECA-Athletic with the Leadership rating (.29); Combined ECA with Military Performance (.26) and MQPR (.28); and Rank in High School Class, Predicted QPR, and Candidate Multiple with MQPR (.29, .33, .37).

Average SAT/ACT Verbal and Mathematical scores, Rank in High School Class, Predicted QPR, and Candidate Multiple correlated moderately to substantially with AQPR (.32, .36, .51, .58, .57,.

For earlier graduates, Average SAT/ACT Verbal scores correlated moderately--and negatively--with the Leadership rating (-.25).

Average SAT/ACT Mathematical score, Rank in High School Class, Predicted QPR, and Candidate Multiple correlated moderately with AQPR (.26, .28, .34, .35).

(Data on the correlations for ECA-Athletic and ECA-Nonathletic are not available for earlier graduates.)

In summary, some of the current predictors, notably ECA-Athletic, had some ability to predict leadership, at least for 1987 graduates. SAT/ACT scores, Rank in High School Class, and composites based on these variables predicted academic performance for both samples. The differential validity of the predictors for the two samples was unexpected.

Conclusions

Validity of Personality Scales

A key finding is that the Sociability scale, and to a lesser extent, the Dominance and Need for Achievement scales, had some validity. The scales demonstrated convergent and discriminant validity in the multitrait-multimethod analyses with peer ratings, and correlated with other relevant predictors and with the leadership criteria. The Sociability scale also correlated with the Conduct and AQPR criteria. However, the level of validity was often modest, particularly for the Dominance and Need for Achievement scales. The failure of the Emotional Stability and Self-Confidence scales to show any sign of validity cannot be explained at this juncture.

The appreciable correlation (.28) between the Sociability scale and the Leadership rating for 1987 graduates, coupled with its small correlations with the secondary criteria of leadership, suggests that

this scale may be useful in selection, at least for new high school graduates. None of the predictors currently used were able to predict this primary criterion of leadership, with the important exception of ECA-Athletic (r = .29) and Combined ECA (r = .24). (The other component of Combined ECA, ECA-Nonathletic, was unrelated to the criterion, r = .00). Although the Sociability scale overlaps with ECA-Athletic (r = .31), it may still make an independent contribution to the prediction of leadership. This possibility is consistent with the observation that the two measures, in combination, had a higher correlation ($p < .01$) with the criterion (R = .35) than either one by itself. However, the correlations for ECA-Athletic may be attenuated because this measure was used in the selection process.

The value of the Sociability scale in selection clearly requires further confirmation. The scale needs to be administered under operational conditions to applicants and its validity--including its incremental validity vis-a-vis ECA-Athletic--appraised in that context. (Minor revisions in the operational version of the Sociability scale are called for because the application process begins during the junior year of high school and some items on the current scale concern the senior year--junior year can be substituted for senior year in these items.) At the same time, the potential trade off between selecting for leadership and for academic ability also needs to be examined. It is noteworthy that the Sociability scale was appreciably and negatively related to SAT/ACT scores (-.28, -.35) for 1987 graduates. ECA-Athletic and Combined ECA had similar

relations with SAT/ACT scores (-.36, -.31 for the former; -.25, -.20 for the latter).

In interpreting the validity results, it must be borne in mind that the ratings and the leadership criteria were less than ideal. The ratings, including the primary leadership criterion, were affected by a halo factor, and the secondary leadership criteria, Military Performance and MQPR, reflect things besides leadership.

The generally modest correlations of the personality scales with the leadership criteria contrast with the generally appreciable correlations of the trait ratings with these criteria. However, it does not necessarily follow that the peer ratings of the traits were more predictive of leadership than were the inventory measures. The correlations between the ratings and the leadership criteria may have been inflated by the same halo factor that affected the intercorrelations of the ratings. The halo factor may have been present in the criteria, just as it was in the trait ratings, the midshipmen's likability intruding not only into the Leadership rating but also into the Military Performance and MQPR grade criteria.

The present results, in total, offer no more than modest support for the proposition that personality traits are implicated in leadership. However, this conclusion needs to be qualified because of the methodological limitations already noted and the specialized nature of the leadership situation being studied: leadership by incoming midshipmen. Whether the present conclusions are generalizable to other contexts in the Academy, in the Navy, or elsewhere is uncertain. As a first step, follow-up studies of the

predictability of leadership in other situations at the Academy would be valuable.

Methodological Implications

Two methodological implications of the results stand out. First, the findings demonstrate that homogeneous scales, made up of biographical items, can be successfully constructed to assess specific traits. The same procedures, adapted from those previously used in developing personality inventories (Jackson, 1970), can readily be applied to the measurement of other kinds of individual-difference variables with biographical items.

Second, the results point to important differences between the 1987 graduates and the earlier graduates. Differences in the validity of the personality scales were anticipated because of divergences in the meaning of their item content for the two groups. However, differences in the correlations of the other predictors with the leadership criteria, as well as with Conduct and AQPR, were surprising.

One dramatic difference between the groups is the negative correlations of the average SAT/ACT Verbal scores with Leadership ratings for earlier graduates and the near zero correlations of these scores for 1987 graduates. It is pertinent that the earlier graduates' SAT/ACT scores were substantially lower (over one-half standard deviation) and their Leadership ratings, as well as their other ratings, were substantially higher than those of 1987 graduates.[6] One conjecture is that, because of their military experience, the earlier graduates who participated in NAPS may be more

liked by their peers and, rightly or wrongly, perceived as having greater leadership ability. The NAPS midshipmen are also apt to have the lowest SAT/ACT scores. The combination of high leadership ratings and low SAT/ACT scores for the NAPS midshipmen, a substantial segment of the earlier graduates, would produce the negative correlation observed between the two variables for these graduates.

Another striking difference is the lower correlations, for earlier graduates, for two combinations of predictors and criteria: (a) Combined ECA with the leadership criteria and (b) Average SAT/ACT scores, Rank in High School Class, and the composites based on these variables (Predicted QPR and Candidate Multiple) with AQPR. The explanation for these lower correlations may center around the greater length of time between when the predictor data were collected for the earlier graduates and when their criterion data were obtained, and the intervening experiences of these midshipmen during that time period. The earlier graduates who participated in NAPS had experiences that might enhance their leadership skills, or at least the perception of their leadership ability by others. And the earlier graduates who participated in NAPS or the U. S. Naval Foundation program, or who attended college, had experiences that could improve their academic performance. Hence, the correlations of the predictors with the criteria for earlier graduates could be attenuated by intervening experiences that differentially affected the criterion performance for subgroups of the earlier graduates.

It would be useful to confirm the reasons for these differences in results for 1987 graduates and earlier graduates. And in future

research on admissions at the Academy, it would be prudent to be alert to the possibility of divergences in the validity and predictability of admissions measures for applicants who are new high school graduates and those who graduated earlier.

References

Aronoff, J., & Wilson, J. P. (1985). Personality in the social process. Hillside, NJ: Erlbaum

Bass, B. M. (1960). Leadership, psychology, and organizational behavior. New York: Harper & Row.

Bass, B. M. (Ed.). (1981). Stogdill's handbook of leadership. New York: Free Press.

Bobko, P., & Rieck, A. (1980). Large sample estimators for standard errors of functions of correlation coefficients. Applied Psychological Measurement, 4, 385-398.

Buss, D. M., & Craik, K. H. (1981). The act frequency analysis of interpersonal dispositions: Aloofness, gregariousness, dominance, and submissiveness. Journal of Personality, 49, 175-211.

Dermen, D., French, J. W., & Harman, H. H. (1978). Guide to factor-referenced temperament tests, 1978. Princeton, NJ: Educational Testing Service.

Gibb, C. A. (1954). Leadership. In G. Lindzey (Ed.), Handbook of social psychology. (Vol. 2, pp. 877-920). Cambridge, MA: Addison-Wesley.

Gibb, C. A. (1969). Leadership. In G. Lindzey & E. Aronson (Eds.), Handbook of social psychology (2nd ed., Vol. 4, pp. 205-282). Reading, MA: Addison-Wesley.

Hogan, R. (1987). The return of the repressed: A theory of social action. [Review of Personality in the social process.] Contemporary Psychology, 32, 43-44.

House, R. J., & Baetz, M. L. (1979). Leadership: Some empirical generalizations and new research directions. In L. L. Cummings & B. M. Shaw (Eds.), Research in organizational behavior (Vol. 1, pp. 341-423). Greenwich, CT: JAI Press.

Jackson, D. N. (1970). A sequential system for personality scale construction. In C. D. Spielberger (Ed.), Current topics in clinical and community psychology. (Vol. 2, pp. 61-96). New York: Academic Press.

Lord, R. G., De Vader, C. L., & Alliger, G. M. (1986). A meta-analysis of the relation between personality traits and leadership preceptions: An application of validity generalization procedures. Journal of Applied Psychology, 71, 402-410.

Mann, R. D. (1959). A review of the relationships between personality and performance in small groups. Psychological Bulletin, 56, 241-271.

Siegel, L. (1958). Manual, the Biographical Inventory for Students (BIS). Princeton, NJ: Educational Testing Service.

Stogdill, R. M. (1948). Personal factors associated with leadership: A survey of the literature. Journal of Psychology, 25, 35-71.

Stricker, L. J. (1987, October). Developing a biographical measure to assess leadership potential. In P. K. Oltman (Chair), Advances in cognitive and personality measurement. Symposium presented at the meeting of the Military Testing Association, Toronto.

Author Note

This research was supported in part by the U. S. Navy-American Society for Engineering Education Summer Faculty Research Program (assembling the biographical inventory) and in part by the Navy Personnel Research and Development Center, U.S. Army Research Office, and Battelle Institute under contract DAAL03-86-D-0001 (item and validity analyses of the inventory, and preparation of this report). The opinions expressed in this report are those of the author, are not official, and do not necessarily reflect the views of the Navy or Army Departments.

Thanks are due Cynthia Benson, Virginia Berry, Rebecca Redard, and Thomas Trent for writing items; Edward F. Alf, Jr., David C. Atwater, John Bruni, Walter Emmerich, Catherine A. Riordan, Robert F. Morrison, George E. Seymour, and Marie D. Thomas for reviewing items; Clayton Culver, Philip K. Oltman, and Carol Staszewski for classifying the items; David M. Buss for providing unpublished act frequency data; Edward F. Alf, Jr., Leanne Atwater, Arthur J. Farkas, Kathryn E. Lesko, and Marie D. Thomas for collecting act frequency data; David C. Atwater, Marie D. Thomas, and Sean Whalen for assistance in pilot testing; Richard L. Deni, Marvin W. Goldstein, Allison McClure, and Robert J. Melara for collecting social desirability data; David C. Atwater, Eric D. Bowman, and Timothy Corrigan for assistance in administering the biographical inventory; David C. Atwater for collecting peer rating data; David C. Atwater, Timothy Corrigan, Richard L. Davis, Steven Devlin, Karen L. LiMato, Joyce D. Mattson, Idell Neumann, and Ann Wahrenbrock for providing data on admissions

measures, grades, and background variables; Philip K. Oltman and Donald A. Rock for advising about the data analysis; Norma A. Norris and Katherine Pashley for computer programming; and Norman M. Abrahams, Robert F. Boldt, Idell Neumann, and Philip K. Oltman for reviewing drafts of this report.

Footnotes

[1]A meta-analysis that appeared subsequently found that dominance and emotional stability were associated with leadership, but sociability was not--need for achievement and self-confidence were not studied (Lord et al., 1986).

[2]This scale was constructed from the 25 Yes-No items on the five tentative personality scales that (a) were not selected for the final scales in a preliminary analysis of the personality scales (the analysis was similar to the one described earlier, with the important exception that no acquiescence measure was included) and (b) had moderate endorsement frequencies (.40 to .60).

An item analysis was carried out, using the same random half of 1987 graduates employed in the item analysis of the personality scales. The 18-item total score was based on two Yes-keyed and two No-keyed items from each personality scale when four such items were available, and all items when fewer were available (two items were randomly chosen when more than two items keyed in the same direction were available).

Biserial correlations were computed between each of the 25 items and the total score. Correlations were corrected for overlap when the item was included in the total score. Items that did not correlate significantly ($p > .10$, one-tail) with the total score were eliminated. Twenty-four items, representing all five personality scales, survived the analysis.

The final scale was based on 17 items from the five personality scales. The items were chosen in the same way that the items were

selected for the total score used in the item analysis, except that when more than two items keyed in the same direction were available for a personality scale, the two with the highest correlations with the total score in the item analysis were chosen.

[3]This scale was constructed from the 36 items on the tentative Social Desirability scale. Multiple-choice items were dichotomized at the median. Items that had extreme endorsement frequencies or had been extensively revised were eliminated. Multiple-choice items were also dropped if the proportion of socially desirable responses, when the items were dichotomized, was not significantly greater than .50 ($p > .05$, two-tail). The 31 surviving items represented 19 personality factors.

An item analysis was carried out. The 19-item total score was based on one item for each factor (when two items were available for a personality factor, a multiple-choice item was chosen over a Yes-No item, and the item with the most extreme social desirability in a pair of Yes-No items was selected).

Biserial correlations were computed between each of the 31 items and the total score. Correlations were corrected for overlap when the item was included in the total score. Items that did not correlate significantly ($p > .10$, one-tail) with the total score were eliminated. Twenty-two items, representing 15 personality factors, survived.

The final scale was based on 12 items, one item for each personality factor. When two items were available for a personality factor, a multiple-choice item was chosen over a Yes-No item, and the Yes-No item with the highest correlation with the total score in the

item analysis in a pair of Yes-No items was selected. In addition, three of the No-keyed with the lowest correlations with the total score in the analysis were excluded to balance the keying.

[4]This scale was constructed from the 31 Yes-No items on the tentative personality scales that (a) were not selected for the final scales in the final item analysis of the personality scales and (b) had moderate endorsement frequencies.

An item analysis was carried out. The 17-item total score was based on items chosen in the same way that the items were selected for the 18-item total score in the item analysis of the initial Acquiescence scale.

Biserial correlations were computed between each of the 31 Yes-No items and (a) this total score and (b) the 19-item total score for the Social Desirability scale used in the item analysis of the second version of the Social Desirability scale (see footnote 4). Correlations with the Acquiescence scale were corrected for overlap when the item was included in the total score, and the correlations with both scales were corrected for attenuation.

An item was eliminated if:

a. It did not correlate significantly ($p > .10$, one-tail) with the total score for the Acquiescence scale (the significance test was applied to the actual correlations).

b. Its correlation with the Acquiescence scale was equaled or exceeded by its correlation with the Social Desirability scale.

Twenty items, representing four personality scales (all except Need for Achievement), survived. The final scale was based on 11

items from the four personality scales. The items were chosen in the same way that items were selected for the final version of the initial Acquiescence scale.

[5]This scale was constructed from the 31 items used in the item analysis of the initial version of the Social Desirability scale. A new item analysis was carried out.

Biserial correlations were computed between each of the 31 items and the 17-item total score for the Acquiescence scale used in the item analysis of the second version of the Social Desirability scale. The correlations with the total score for the Acquiescence scale and the previously computed correlations with the 19-item total score used in the item analysis of the initial version of the Social Desirability scale were corrected for attenuation.

An item was eliminated if:

a. It did not correlate significantly with the total score for the Social Desirability scale.

b. Its correlation with the Social Desirability scale was equaled or exceeded by its correlation with the Acquiescence scale.

Thirteen items, representing 12 personality factors, survived. The final scale was based on 12 of these items, one for each factor. When two items were available for a personality factor, a multiple-choice item was chosen over a Yes-No item.

[6]The corresponding means (and standard deviations) for 1987 graduates used in the validity analysis and earlier graduates were 564.13 (70.35) and 523.51 (69.52) for Average SAT/ACT Verbal score, 642.34 (64.19) a 594.61 (69.01) for Average SAT/ACT Mathematical

score, and 48.13 (9.59) and 55.16 (8.61) for Leadership rating. The two samples' means (and standard deviations) for the other ratings were 48.22 (9.94) and 54.70 (8.55) for Dominance, 48.30 (9.60) and 54.48 (7.79) for Emotional Stability, 48.57 (9.61) and 53.05 (9.56) for Need for Achievement, 48.64 (9.69) and 54.41 (7.87) for Self-Confidence, and 49.20 (9.86) and 53.39 (8.98) for Sociability.

Table 1

Summary of Reviews of Personality Correlates of Leadership

Personality Trait	Review						
	Aronoff & Wilson (1985)	Bass (1960)	Gibb (1954)	Gibb (1969)	House & Baetz (1979)	Mann (1959)	Stogdill (1948)
Authoritarianism	Yes					Yes	
Dominance	Yes		No	No	Yes	Yes	No
Emotional Stability	Yes					Yes	No
Empathy		NA		NA		NA	Yes
Energy		Yes					Yes
Need for Achievement	Yes	Yes					Yes
Self-Confidence	Yes		Yes	Yes			Yes
Sociability	Yes		No			Yes	Yes
Surgency	Yes		Yes	No		Yes	No

Note. Yes = trait is generally related to leadership in studies reviewed, No = trait is not generally related to leadership, NA = relationship cannot be ascertained because of methodological flaws in studies.

Table 2

Definitions of Personality Traits

Dominance
1. Takes charge socially, wants power vs. submissive, willing to serve.
2. Egoistic, pushes own ideas vs. respects others' ideas, self-effacing.
3. Rights-conscious, complaining vs. tolerant.

Emotional Stability
1. Emotionally stable, tolerant, stolid vs. emotionally sensitive, irritable.
2. Optimistic, faces problems vs. worrying, dwells on problems, escapist.
3. Feels healthy vs. hypochondriacal.

Need for Achievement
1. Likes success in competition, likes getting ahead vs. dislikes competition.
2. Strives for accomplishment, wants to produce something great.

Self-Confidence
1. Feels confident physically, personally, and career-wise vs. needs encouragement, feels inferior, afraid of failure.
2. Claims to have abilities, skills, and good experiences vs. claims handicaps, ineptitude, and unfavorable experiences.
3. Perceives others as having been positive toward him vs. negative.

Sociability
1. Glib talker, has superficial social know-how vs. aloof, doesn't know or care what should be said.
2. Hardened socially, confident in social contacts vs. shy, socially insecure.
3. Competent socially, social organizer, enjoys attention vs. withdrawn, fears public speaking and social responsibilities.

Source: Dermen, Harman, and French (1978).

Table 3

Number of Items on Personality Scales at Each Step of Item Analysis

Scale	Administered			First Item Analysis			Second Item Analysis			Final Scale
	Total	Substantial Agreement	Borderline Agreement	Total Score	Other	Surviving	Total Score	Other	Surviving	
Dominance	45	23	22	23	22	14	11	34	14	14
Emotional Stability	43	23	15	28	15	14	14	29	17	17
Need for Achievement	42	29	13	21	21	21	21	21	17	17
Self-Confidence	60	32	28	32	28	18	8	52	16	8
Sociability	50	40	10	38	12	19	19	31	22	22

Table 4

Intercorrelations of Personality and Response Style Scales, Sex, and Ethnicity for 1987 Graduates

Variable	1	2	3	4	5	6	7	8	9
1. Dominance scale	(.68)	-.12	.39	-.09	.57	.19	.22	-.02	-.04
2. Emotional Stability scale		(.68)	.04	.23	-.06	-.25	.28	-.01	-.08
3. Need for Achievement scale			(.66)	-.22	.44	.09	.36	-.06	-.10
4. Self-Confidence scale				(.67)	-.16	-.34	.14	-.06	.15
5. Sociability scale					(.78)	.10	.37	.04	-.03
6. Acquiescence scale						(.58)	-.27	-.02	.05
7. Social Desirability scale							(.27)	-.14	-.11
8. Sex								(--)	.03
9. Ethnicity									(--)

Note. Coefficient Alpha reliability coefficients appear in parentheses. N is 233. Correlations of .13 and .17 are significant at the .05 and .01 levels (two-tail), respectively.

Table 5

Intercorrelations of Personality and Response Style Scales, Sex, and Ethnicity for Earlier Graduates

Variable	1	2	3	4	5	6	7	8	9
1. Dominance scale	(.62)	.11	.36	-.13	.51	.14	.11	.09	.02
2. Emotional Stability scale		(.69)	.11	.09	.20	-.26	.35	-.06	-.06
3. Need for Achievement scale			(.64)	-.27	.28	-.04	.35	-.04	-.04
4. Self-Confidence scale				(.64)	-.14	-.28	.01	.04	.14
5. Sociability scale					(.77)	.05	.28	.03	.03
6. Acquiescence scale						(.57)	-.19	.08	-.01
7. Social Desirability scale							(.35)	-.06	.03
8. Sex								(--)	.08
9. Ethnicity									(--)

Note. Coefficient Alpha reliability coefficients appear in parentheses. N is 171. Correlations of .15 and .20 are significant at the .05 and .01 levels (two-tail), respectively.

Table 6

Intercorrelations of Criteria for 1987 Graduates

Variable	1	2	3	4	5	6	7	8	9	10
1. Dominance rating	(.87)	.63	.59	.84	.78	.82	.39	.37	.16	.16
2. Emotional Stability rating		(.76)	.57	.74	.44	.73	.41	.41	.19	.19
3. Need for Achievement rating			(.86)	.64	.45	.70	.52	.51	.22	.26
4. Self-Confidence rating				(.86)	.71	.82	.38	.37	.16	.19
5. Sociability rating					(.88)	.71	.26	.16	.01	-.02
6. Leadership rating						(.89)	.44	.39	.18	.16
7. Military Performance							(--)	.85	.31	.42
8. MQPR								(--)	.52	.59
9. Conduct									(--)	.20
10. AQPR										(--)

Note. Interrater reliability coefficients appear in parentheses. Ns vary from 193 to 223. For an N of 223, correlations of .13 and .17 are significant at the .05 and .01 levels (two-tail), respectively.

Table 7

Intercorrelations of Criteria for Earlier Graduates

Variable	1	2	3	4	5	6	7	8	9	10
1. Dominance rating	(.87)	.53	.63	.82	.79	.78	.37	.33	.10	.03
2. Emotional Stability rating		(.83)	.54	.78	.36	.76	.23	.23	-.03	.11
3. Need for Achievement rating			(.81)	.68	.39	.72	.41	.42	.17	.23
4. Self-Confidence rating				(.84)	.66	.86	.32	.29	.06	.11
5. Sociability rating					(.87)	.59	.17	.10	.02	-.10
6. Leadership rating						(.87)	.32	.29	.07	.08
7. Military Performance							(--)	.89	.14	.32
8. MQPR								(--)	.25	.47
9. Conduct									(--)	.08
10. AQPR										(--)

Note. Interrater reliability coefficients appear in parentheses. Ns vary from 153 to 166. For an N of 166, correlations of .15 and .20 are significant at the .05 and .01 levels (two-tail), respectively.

Table 8

Intercorrelations of Peer Ratings for Random Halves of Raters for 1987 Graduates

Rating	1	2	3	4	5	6	7	8	9	10	11	12
First Half of Raters												
1. Dominance		.58	.57	.79	.73	.75	.78	.47	.56	.73	.71	.72
2. Emotional Stability			.51	.73	.47	.71	.56	.62	.52	.63	.45	.63
3. Need for Achievement				.59	.44	.64	.51	.41	.76	.49	.41	.60
4. Self-Confidence					.65	.77	.75	.60	.65	.76	.65	.78
5. Sociability						.67	.67	.26	.39	.60	.79	.58
6. Leadership							.73	.55	.65	.69	.63	.79
Second Half of Raters												
7. Dominance								.55	.59	.81	.75	.81
8 Emotional Stability									.51	.60	.30	.60
9. Need for Achievement										.64	.45	.73
10. Self-Confidence											.70	.75
11. Sociability												.66
12. Leadership												

Note. Ns vary from 218 to 223. For an N of 223, correlations of .13 and .17 are significant at the .05 and .01 levels (two-tail), respectively

Table 9

Intercorrelations of Peer Ratings for Random Halves of Raters for Earlier Graduates

Rating	1	2	3	4	5	6	7	8	9	10	11	12
First Half of Raters												
1. Dominance		.52	.54	.80	.69	.73	.77	.52	.52	.68	.69	.66
2. Emotional Stability			.49	.72	.31	.67	.41	.72	.43	.62	.33	.55
3. Need for Achievement				.56	.30	.65	.51	.44	.68	.55	.31	.57
4. Self-Confidence					.61	.80	.63	.66	.52	.72	.60	.70
5. Sociability						.52	.66	.26	.28	.47	.78	.48
6. Leadership							.62	.66	.58	.66	.49	.77
Second Half of Raters												
7. Dominance								.51	.59	.75	.76	.74
8. Emotional Stability									.55	.73	.38	.74
9. Need for Achievement										.68	.45	.71
10. Self-Confidence											.63	.80
11. Sociability												.61
12. Leadership												

Note. Ns vary from 164 to 166. For an N of 166, correlations of .15 and .20 are significant at the .05 and .01 levels (two-tail), respectively.

Table 10

Correlations of Personality Scales with Criteria for 1987 Graduates

Personality Scale	Rating									
	Domi-nance	Emotional Stability	Need for Achievement	Self-Confidence	Socia-bility	Leader-ship	Military Performance	MQPR	Conduct	AQPR
Dominance	.23	.01	.06	.13	.20	.06	.01	.04	.01	-.04
Emotional Stability	-.11	.03	.04	-.09	-.19	.06	-.02	.02	.03	.00
Need for Achievement	.08	-.01	.12	.03	.07	.07	.12	.13	.09	.08
Self-Confidence	.10	.14	.15	.12	.03	.09	.03	.07	-.02	.06
Sociability	.35	.13	.21	.27	.36	.28	.13	.12	.16	-.04

Note. Ns vary from 199 to 223. For an N of 223, correlations of .13 and .17 are significant at the .05 and .01 levels (two-tail), respectively.

Table 11

Correlations of Personality Scales with Criteria for Earlier Graduates

Personality Scale	Rating									
	Dominance	Emotional Stability	Need for Achievement	Self-Confidence	Sociability	Leadership	Military Performance	MQPR	Conduct	AQPR
Dominance	.20	.11	.08	.15	.25	.14	.11	.09	-.01	.05
Emotional Stability	-.14	-.05	-.05	-.07	-.01	-.13	-.13	-.09	-.06	.01
Need for Achievement	.08	-.03	.11	.06	.14	.00	.07	.10	.03	.06
Self-Confidence	-.03	-.08	.02	-.09	-.10	-.07	-.03	.02	.05	.02
Sociability	.27	.15	.07	.26	.39	.16	-.04	-.09	-.04	-.17

Note. Ns vary from 156 to 166. For an N of 166, correlations of .15 and .20 are significant at the .05 and .01 levels (two-tail), respectively.

Table 12

Correlations of Personality Scales with Other Predictors for 1987 Graduates

Predictor	Personality Scale				
	Dominance	Emotional Stability	Need for Achievement	Self-Confidence	Sociability
Average SAT/ACT Verbal score	-.14	.08	-.11	.09	-.28
Average SAT/ACT Mathematical score	-.21	.09	-.07	.13	-.35
Rank in high school class	.00	.04	.17	.08	-.08
Recommendations of high school officials	.09	-.02	.07	.01	.09
ECA-Athletic	.13	-.14	.02	.07	.31
ECA-Nonathletic	.39	.11	.17	.02	.32
Combined ECA	.39	-.06	.17	.05	.47
Predicted QPR	-.11	.08	.06	.12	-.25
Candidate Multiple	-.04	.10	.11	.10	-.15

Note. N is 233. Correlations of .13 and .17 are significant at the .05 and .01 levels (two-tail), respectively.

Table 13

Correlations of Personality Scales with Other Predictors for Earlier Graduates

Predictor	Personality Scale				
	Dominance	Emotional Stability	Need for Achievement	Self-Confidence	Sociability
Average SAT/ACT Verbal score	.02	.06	-.09	.21	-.17
Average SAT/ACT Mathematical score	.06	-.05	-.02	.03	-.08
Rank in high school class	.02	.02	.20	.13	-.19
Recommendations of high school officials	.02	-.01	-.10	-.02	-.04
ECA-Athletic[a]	--	--	--	--	--
ECA-Nonathletic[a]	--	--	--	--	--
Combined ECA	.29	.17	.14	.11	.23
Predicted QPR	.04	.02	.09	.17	-.22
Candidate Multiple	.11	.03	.09	.14	-.17

Note. N is 171. Correlations of .15 and .20 are significant at the .05 and .01 levels (two-tail), respectively.

[a]Correlations for this predictor are not available.

Table 14

Correlations of Other Predictors with Criteria for 1987 Graduates

Predictor	Criteria				
	Leadership Rating	Military Performance	MQPR	Conduct	AQPR
Average SAT/ACT Verbal score	-.01	.06	.16	-.02	.32
Average SAT/ACT Mathematical score	-.03	.10	.23	.04	.36
Rank in high school class	.00	.17	.29	.16	.51
Recommendations of high school officials	-.04	.14	.12	-.05	.15
ECA-Athletic	.29	.22	.20	.16	.05
ECA-Nonathletic	.00	.11	.14	06	.07
Combined ECA	.24	.26	.28	.17	11
Predicted QPR	-.01	.17	.33	.12	.58
Candidate Multiple	.01	.24	.37	.13	.57

Note. Ns vary from 199 to 218. For an N of 218, correlations of .13 and .17 are significant at the .05 and .01 levels (two-tail), respectively.

Table 15

Correlations of Other Predictors with Criteria for Earlier Graduates

Predictor	Criteria				
	Leadership Rating	Military Performance	MQRP	Conduct	AQPR
Average SAT/ACT Verbal score	-.25	.07	.14	-.06	.14
Average SAT/ACT Mathematical score	-.16	.09	.16	-.10	.26
Rank in high-school class	-.07	.03	.07	-.07	.28
Recommendations of high school officials	.12	.02	.09	.03	.02
ECA athletic[a]	--	--	--	--	--
ECA non-athletic[a]	--	--	--	--	--
Combined ECA	.08	.05	-.01	-.02	-.01
Predicted QPR	-.19	.08	.16	-.07	.34
Candidate Multiple	-.11	.14	.20	-.04	.35

Note. Ns vary from 156 to 164. For an N of 164, correlations of .15 and .20 are significant at the .05 and .01 levels (two-tail), respectively.

[a]Correlations for this predictor are not available.

Figure Caption

Figure 1. Rating scales for peer ratings. (Scores used to quantify ratings are in brackets.)

Dominance

	Extremely Dominant	Very Dominant	Somewhat Dominant	Slightly Dominant	Slightly Submissive	Somewhat Submissive	Very Submissive	Extremely Submissive	
DOMINANT: Takes charge socially, pushes own ideas and interests, stands up for rights.	1 [9]	2 [8]	3 [7]	4 [6]	5 [4]	6 [3]	7 [2]	8 [1]	SUBMISSIVE: Defers to others, not forward, does not complain.

Emotional Stability

	Extremely Emotionally Stable	Very Emotionally Stable	Somewhat Emotionally Stable	Slightly Emotionally Stable	Somewhat Affected by Feelings	Somewhat Affected by Feelings	Very Affected by Feelings	Extremely Affected by Feelings	
EMOTIONALLY STABLE: Unemotional, faces problems, feels healthy.	1 [9]	2 [8]	3 [7]	4 [6]	5 [3]	6 [3]	7 [2]	8 [1]	AFFECTED BY FEELINGS: Easily upset, worries, hypochondriacal.

Need for Achievement

	Extremely Motivated to Achieve	Very Motivated to Achieve	Somewhat Motivated to Achieve	Slightly Motivated to Achieve	Slightly Unambitious	Somewhat Unambitious	Very Unambitious	Extremely Unambitious	
MOTIVATED TO ACHIEVE: Wants to get ahead, strives for accomplishment.	1 [9]	2 [8]	3 [7]	4 [6]	5 [4]	6 [3]	7 [2]	8 [1]	UNAMBITIOUS. Dislikes competing.

Self-Confidence

	Extremely Self-Conscious	Very Self-Conscious	Somewhat Self-Conscious	Slightly Self-Conscious	Slightly Self-Assured	Somewhat Self-Assured	Very Self-Assured	Extremely Self-Assured	
SELF-CONSCIOUS: Needs encouragement, thinks he or she is incapable, believes he or she is not liked and respected.	1 [9]	2 [8]	3 [7]	4 [6]	5 [5]	6 [4]	7 [3]	8 [2]	SELF-ASSURED: Feels confident, thinks he or she is capable, believes he or she is liked and respected.

Sociability

	Extremely Shy	Very Shy	Somewhat Shy	Slightly Shy	Slightly Sociable	Somewhat Sociable	Very Sociable	Extremely Sociable	
SHY: Distant in relations with others, socially insecure, socially retiring.	1 [9]	2 [8]	3 [7]	4 [6]	5 [5]	6 [4]	7 [3]	8 [2]	SOCIABLE: Has social know-how, socially confident, enjoys being the center of attention.

Leadership

	Extremely Able to Lead	Very Able to Lead	Somewhat Able to Lead	Slightly Able to Lead	Slightly Unable to Lead	Somewhat Unable to Lead	Very Unable to Lead	Extremely Unable to Lead	
ABLE TO LEAD: Directs a group's activities, obtains a group's cooperation.	1 [9]	2 [8]	3 [7]	4 [6]	5 [5]	6 [4]	7 [3]	8 [2]	UNABLE TO LEAD: Does not direct a group's activities, does not obtain a group's cooperation.