Reproduced by the

ARMED SERVICES TECHNICAL INFORMATION AGENCY
ARLINGTON HALL STATION
ARLINGTON 12, VIRGINIA
NOTICE: When government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related government procurement operation, the U. S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.
OFFICER PEER RATINGS AS A PREDICTOR OF FAILURE TO COMPLETE FLIGHT TRAINING

SPECIAL REPORT NO. 62-2

By

Lieutenant (jg) Richard E. Doll, MSC, USNR

Approved by

Captain Ashton Graybiel, MC, USN
Director of Research

Released by

Captain Clifford P. Phoebus, MC, USN
Commanding Officer

25 April 1962

Opinions or conclusions contained in this report are those of the author. They are not to be construed as necessarily reflecting the view or the endorsement of the Navy Department. Reference may be made to this report in the same way as to published articles noting author, title, source, date, and report number.
THE PROBLEM

Peer ratings have proved to be useful instruments in appraising the potential of individual cadets in the Naval Air Training Program. Almost half of the students in naval air training, however, are officers rather than cadets. A valid rating of the training potential of these would also be useful. There is a considerable difference, however, between the cadet and officer situation during the early stage of training; in contrast to the cadets' "around-the-clock" mutual association, the officer students are associated as a group only during academic and physical training instruction.

This study investigated the predictive validity of peer ratings administered to 28 classes of officer students. The predictive validity of these ratings is compared with that of similar ratings executed by 28 classes of cadets. Multiple correlation procedures were used to determine the extent to which the peer ratings made an independent contribution to the predictive validity of the selection test scores and training grades.

FINDINGS

It was found that while the cadet peer ratings made a substantial contribution to the multiple validity for predicting success, the officer peer ratings did not.

CONCLUSION

It was concluded that the introduction of officer peer ratings as a regular evaluation device in naval aviation training would not be justified by the findings of the study.
INTRODUCTION

Peer ratings, i.e., evaluations of the members of a group by one or more other members of that group, have proved to be useful instruments. Although usually made by untrained and relatively unsophisticated observers, such ratings have been shown to be good predictors of success or failure in several areas of endeavor.

Williams and Leavitt (6) investigated peer nominations as predictors of Officer Candidate School (OCS) grades and leadership performance in combat. For both of these behaviors peer ratings were found to be superior to several objective tests and to ratings by superiors. The authors attributed this to the fact that "group members have more time to observe each other than do superior officers, they know each other in a realistic social context, and they react directly to each other's social dominance behavior."

La Gaipa (3) found that none of the OCS selection variables used at Newport predicted on-the-job success as measured by fitness reports for officers assigned to either shore or fleet billets. However, a significant correlation of .23 resulted between peer ratings administered in the eleventh week of OCS training and the later fitness reports of both shore and fleet assigned officers.

Baier (1) has shown that peer nominations at West Point were more highly related to later performance as officers than was scholastic standing; and Rigby and Ossorio (5) reported that peer ratings administered early in the training of Naval and Marine Corps women recruits discriminated between successful and nonsuccessful trainees.

Since 1957 such peer ratings have been among the measures used in the Naval Air Training Program to appraise the potential of individual cadets. These ratings are administered during the eighth week of pre-flight school training. This training situation is quite similar to that at OCS, Newport, inasmuch as the cadets are divided into classes and attend all phases of military, academic, and physical training as a class. In addition, they eat together and live in the same quarters. The two situations are also similar in that both are of sixteen weeks' duration and upon completion cadets with college degrees are commissioned. The peer rating format used in naval air training calls for each man in a class of cadets to name the three most promising prospective officers and the three least promising in his class. It has been shown (2,4) that these ratings, when pooled, typically have bi-serial correlations of about .35 to .40 with subsequent failure to complete the training program, and that, when combined with other measures, they have considerable administrative usefulness.
Almost half of the students in naval air training, however, are officers rather than cadets. A valid rating of the training potential of these men would also be administratively useful. There is a considerable difference, however, between the cadet and the officer situations at this stage of training. In contrast to the cadets' "around-the-clock" mutual association, the officer students as a class go only to academic and physical training instruction together. After the working day they go their own separate ways, either home to their families, or to some type of bachelor quarters.

Since the officer student is with most of his peers only during periods of training, the question that needed answering was, "Do the officer students gain as good insight into their peers' potential as do cadets, without the cadets' opportunity to observe 'around-the-clock' behavior?"

**PROCEDURE**

From June 1960 to July 1961 peer ratings were obtained from 28 officer student classes (N = 606) who were in their tenth week of pre-flight training. These ratings were compared with later success in flight training as indicated by whether or not the officer was still in the program at the end of basic training (about ten months later). All but about 4 per cent of the failures and withdrawals customarily take place before this point in training. The bi-serial correlation coefficient was computed. For comparative purposes the same computation was made of the validity of peer ratings for 28 cadet classes (N = 666) who went through pre-flight school during 1959.

Since a new measurement procedure is useful only if it adds to the predictive validity of other measurements already in use, multiple correlation analysis seemed appropriate. Available for this purpose were scores from three aptitude tests -- the Spatial Apperception Test (SAT), the Mechanical Comprehension Test (MCT), and the Aviation Qualification Test (AGT); the Biographical Inventory (BI); four academic grades -- aviation physiology (AP), engines (ENG), aerodynamics (AERO), and navigation (NAV); and a physical training (PT) grade.

Two multiple correlations were computed for both the officer and cadet groups. One of the multiple correlations for each group included the peer rating score; the other did not. Comparison of the two correlations indicates the peer rating's value as an added predictor of failure to complete flight training.
RESULTS AND DISCUSSION

The bi-serial correlations of the peer ratings with their respective criteria were .20 for the officer students and .36 for the cadets. The cadet peer rating is by far the more valid predictor. However, a correlation of .20 has value if it adds any unique variance to a battery of predictors.

The results expressed in Table I unequivocally deny any such potential. The multiple correlation for the officer student group reaches a maximum of .49, with the peer rating contributing nothing to the multiple. In contrast, the cadet peer rating is second only to the navigation grade in contribution to the multiple. The cadet multiple correlation which includes peer rating scores is .56; when the same multiple is computed omitting the peer rating, it is only .51.

Table I

Comparison of the Multiple Correlations of Officers' and Cadets' Selection, Training, and Peer Rating Scores with a "Failure to Complete" Criterion

<table>
<thead>
<tr>
<th>Variable</th>
<th>Officers</th>
<th>Multiple R*</th>
<th>Cadets</th>
<th>Variable</th>
<th>Multiple R</th>
</tr>
</thead>
<tbody>
<tr>
<td>AERO</td>
<td>.37</td>
<td></td>
<td>NAV</td>
<td>.39</td>
<td></td>
</tr>
<tr>
<td>MCT</td>
<td>.42</td>
<td></td>
<td>P-R</td>
<td>.48</td>
<td></td>
</tr>
<tr>
<td>PT</td>
<td>.45</td>
<td></td>
<td>ENG</td>
<td>.51</td>
<td></td>
</tr>
<tr>
<td>SAT</td>
<td>.47</td>
<td></td>
<td>SAT</td>
<td>.53</td>
<td></td>
</tr>
<tr>
<td>BI</td>
<td>.48</td>
<td></td>
<td>PT</td>
<td>.54</td>
<td></td>
</tr>
<tr>
<td>NAV</td>
<td>.49</td>
<td></td>
<td>BI</td>
<td>.55</td>
<td></td>
</tr>
<tr>
<td>AQT</td>
<td>.49</td>
<td></td>
<td>AERO</td>
<td>.55</td>
<td></td>
</tr>
<tr>
<td>ENG</td>
<td>.49</td>
<td></td>
<td>AP</td>
<td>.56</td>
<td></td>
</tr>
</tbody>
</table>

* Determined by Wherry-Doolittle procedure

The intercorrelations shown in Table II partially explain these results. The peer rating scores of the officer students correlate consistently higher with the training grades than do the cadet peer rating scores. The officer students apparently were basing their judgments of their peers more on classroom behavior than were the cadets.
Table II

Intercorrelation of Variables for Officer Student and Cadet Groups*

<table>
<thead>
<tr>
<th></th>
<th>P-R</th>
<th>AP</th>
<th>ENG</th>
<th>AERO</th>
<th>PT</th>
<th>NAV</th>
<th>SAT</th>
<th>BI</th>
<th>MCT</th>
<th>AQT</th>
</tr>
</thead>
<tbody>
<tr>
<td>P-R</td>
<td>.23</td>
<td>.36</td>
<td>.30</td>
<td>.24</td>
<td>.28</td>
<td>.05</td>
<td>.08</td>
<td>.19</td>
<td>.11</td>
<td></td>
</tr>
<tr>
<td>AP</td>
<td>.11</td>
<td>.12</td>
<td>.19</td>
<td>.23</td>
<td>.25</td>
<td>.12</td>
<td>.12</td>
<td>.08</td>
<td>.13</td>
<td></td>
</tr>
<tr>
<td>ENG</td>
<td>.55</td>
<td>.45</td>
<td>.08</td>
<td>.33</td>
<td>.12</td>
<td>.10</td>
<td>.38</td>
<td>.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AERO</td>
<td>.41</td>
<td>.34</td>
<td>.00</td>
<td>.24</td>
<td>.05</td>
<td>.07</td>
<td>.24</td>
<td>.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PT</td>
<td>.15</td>
<td>.00</td>
<td>.12</td>
<td>.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NAV</td>
<td></td>
<td>.16</td>
<td>.10</td>
<td>.33</td>
<td>.48</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAT</td>
<td></td>
<td></td>
<td>.04</td>
<td>.25</td>
<td>.21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BI</td>
<td></td>
<td></td>
<td></td>
<td>-.05</td>
<td>.12</td>
<td>.16</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AQT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.28</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* The upper correlation is for the officer student group while the lower correlation is that of the cadet group.
Thus the officer peer ratings contributed nothing new to the predictive validity of the training and selection scores. Cadet peer ratings, on the other hand, were apparently based on observed aspects of behavior which were important to later success in training and which were not so highly reflected by the other available measures.

The difference between the contributions of the cadet and officer peer ratings can be most easily attributed to their differences in opportunity to observe their peers under a variety of circumstances. It must be concluded that the introduction of officer peer ratings as a regular evaluation device in naval aviation training would not be justified by the findings of this study.

REFERENCES


