File

DEPARTMENT OF PSYCHOLOGY ST. LOUIS UNIVERSITY

2 additional cys req
9/2/53

11/15/54 1cy John F. Scott Chief
Test Development Section
US C.S. Commission
Wash. 25, D.C.

PEER RATINGS

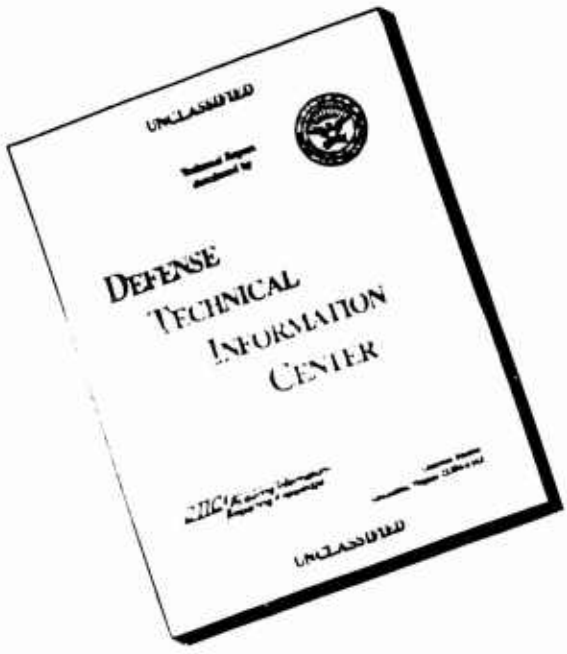
Relationships Between Officer and Peer-candidate Predictions of Effectiveness as a Company Grade Officer in the U S Marine Corps And the Ability to Predict Estimated Officer Effectiveness of Peers

O. F. Anderhalter
Walter L. Wilkins
Marilyn K. Rigby

Technical Report No. 2
November 30, 1952

Prepared for the Neuropsychiatry Branch, Professional Division Bureau of Medicine and Surgery, Department of the Navy, Under Office of Naval Research Contract N7onr - 40802 (NR 151 - 092)

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

DEPARTMENT OF PSYCHOLOGY ST. LOUIS UNIVERSITY

PEER RATINGS

Relationships Between Officer and Peer-candidate Predictions Of Effectiveness as a Company Grade Officer in the U S Marine Corps And the Ability to Predict Estimated Officer Effectiveness of Peers

O. F. Anderhalter
Walter L. Wilkins
Marilyn K. Rigby

Technical Report No. 2
November 30, 1952

Prepared for the Neuropsychiatry Branch, Professional Division Bureau of Medicine and Surgery, Department of the Navy, Under Office of Naval Research Contract N7onr - 40802 (NR 151 - 092)

TABLE OF CONTENTS

		Page
I.	Introduction.	1
II.	Candidates as Perceivers of Officer Potential	4
III.	Similarity of Judgments of Officer Potential	17
IV.	Relationship Between Peer Rating Obtained and Line Officer Rank.	24
V.	Summary and Conclusions	26

I. INTRODUCTION

Military researches of World War II have suggested that peer rating techniques may be usefully employed as evaluation instruments. Wherry and Fryer (7) have reported promising results with a peer nominating technique using samples of Signal Corps officer candidates. They found substantial reliability of the instrument over a period of three months; the reliability coefficients obtained from the ratings of peers greatly exceeded those obtained from supervisors after periods of one and three months. Their data also suggested that peer ratings after one month measured the same things that supervisors' ratings did after four months of observation. McClure, Tupes, and Dailey (3) report work with peer ratings applied to USAF Officer Candidate School members. They found good reliability of the peer rankings as well as a high relationship between those ratings and later on-the-job proficiency ratings. Most pertinent here, perhaps, is the work of Williams and Leavitt (5, 6) with a group of Marine Corps officer candidates. They asked men at the Pre-officer Candidate School at Camp Lejeune to nominate others in their platoon whom they felt would be the best and least desirable

all-round officers. A biserial correlation of about .40 was reported between these nominations and Officer Candidate School standing.

Peer ratings may be analyzed from two points of view. Conventionally, a measure of effectiveness of a man is taken as the consensus of his peers' opinion of him. Cronbach has suggested that analysis of the rankings that a candidate makes of the members of his group might also be relevant to some criteria. Thus, the ability of a man to recognize officer potential in his peers might be an indication of his own potential as an officer. In addition to studying a man as he is perceived by the group, then, it is also possible to study him as he perceives that group. The present study will report findings based on both kinds of approaches to peer ratings.

The U. S. Marine Corps has established a continuing officer screening course at Marine Corps Schools, Quantico, Virginia, where selected enlisted men are examined to estimate their potential effectiveness as officers. The men are selected on the basis of several combinations of Army General Classification Test scores, college education, and commanding officer recommendation. The course itself is described in more complete detail in an earlier publication, The Potential Combat Officer (4).

Data from a group of sixty candidates screened in May, 1952, were used in the present study. The men had been divided into six approximately equal sections, who lived and worked together during the three weeks' course. Every week of the course each man was required to rank-order the other men in his section according to overall officer effectiveness. Peer ratings obtained during the first and third weeks served as primary data for this analysis.

At the end of the screening course Marine Corps assessment staff arranged all sixty men in an overall rank order. This final rank order was based on pooled judgments of performance in a wide variety of situations relevant to the requirements of successful junior officers in the Corps (4), and provided a criterion measure for the study.

II. CANDIDATES AS PERCEIVERS OF OFFICER POTENTIAL

This analysis is concerned with the relationship between the ability of a candidate to evaluate officer potential in his peers and his own adjudged potential as an officer in the Corps. The criterion information used was the final standing of the individual at the end of the screening course based on the opinion of the assessment staff at Quantico after an opportunity to observe the men for three weeks in a wide variety of situations. A criterion of performance in the field would, of course, be more desirable to use in this situation; unfortunately such a criterion was not available for this group. The criterion used, nevertheless, has certain advantages beyond availability. The line officers evaluating the men during the assessment period had been trained for this purpose and were experienced in dealing with men in practical field situations. In addition the recommendations of the assessment staff are utilized as the principal basis in selecting men for commissioning.

A. Selection of Comparison Index

The line officer evaluation at the end of the course orders all sixty men in one list, whereas the ratings made by the candidates

involve only the men in their section. Before an analysis of the relationship between peer ratings and line officer rankings could be effected, the two measures had to be made directly comparable. This was accomplished by listing the assessment rankings of the men in a particular section and reassigning these ranks on the basis of 1 to the number of men in the section, maintaining the same relative position as in the total list. Thus, for the men in the first section the line officer ranks of 3, 13, 16, 18, 28, 30, 39, 41, 49, 51, and 67 were reassigned values of 1 through 11. The original rank and the reassigned rank are presented in Table I. Since each candidate omitted himself in rank ordering the members of his section, it was necessary to reassign the line officer ranks once again for the correlations to be computed. To illustrate, the men ranked by Candidate 1a, who had been assigned the rank of 3 by the officers, would have the line officer ranking of 1, 2, 4, 5, 6, 7, 8, 9, 10, and 11, omitting the rank of 3, since 1a did not rank himself. For the evaluation of Candidate 1a's rankings of his peers therefore, ranks of 3 through 10 were assigned to the candidates formerly having line officer ranks of 4 through 11. Similar reassignment of ranks was made for each individual.

TABLE I

Original and Reassigned Line Officer Ranks

Candidate	Original Rank	Reassigned Rank
Section 1		
1a	16	3
1b	18	4
1c	3	1
1d	13	2
1e	57	11
1f	41	8
1g	49	9
1h	51	10
1i	30	6
1j	28	5
1k	39	7
Section 2		
2a	24	5
2b	20	3
2c	53	9
2d	48	8
2e	55	10
2f	22	4
2g	15	2
2h	59	11
2i	1	1
2j	45	6
2k	46	7
Section 3		
3a	31	6
3b	9	3
3c	5	1
3d	7	2
3e	58	11
3f	11	4
3g	35	9.5
3h	43	9.5
3i	33	7
3j	37	8
3k	26	5
Section 4		
4a	27	3
4b	21	2
4c	47	7
4d	8	1
4e	60	9
4f	54	8
4g	34	4
4h	38	6
4i	36	5
Section 5		
5a	32	5
5b	2	1
5c	40	6
5d	29	4
5e	42	7
5f	10	3
5g	44	8
5h	6	2
Section 6		
6a	52	9
6b	19	5
6c	17	4
6d	23	6
6e	50	8
6f	25	7
6g	12	2
6h	4	1
6i	56	10
6j	14	3

B. Judgment Index and Line Officer Rank

Spearman rank-difference correlation coefficients were computed between the ranking done by each candidate of the other men in the section and the ranking decided upon by the line officers. For purposes of this report, the coefficient representing the relationship between an individual candidate's opinion of how his peers should be ranked and the line officers' opinion will be called the "judgment index." The judgment indices themselves were ranked within each section and that rank is recorded with the indices in Table II. In order to determine the extent to which the judgment index of a candidate is related to his adjudged potential, it was planned to correlate the indices with the final standing of each man on the line officer list. This, however, is not as clear-cut a procedure as it might at first appear. For one thing, no two correlations would be exactly comparable, even for men in the same section, since none of the men in the section rated exactly the same set of individuals, having omitted themselves in their rankings. This limitation is negligible in larger sections, but could be a serious limiting factor in those sections with smaller numbers of men. Another consideration, certainly, is the comparability of the rating task for the various sections. For

TABLE II

Judgment Indices for First and Third Week Rankings

Candidate	First Period Index	First Period Rank	Third Period Index	Third Period Rank
Section 1				
1a	.430	4	.624	4
1b	-.018	10	.418	8
1c	.140	6.5	.649	3
1d	.406	5	.419	7
1e	.043	9	-.357	11
1f	.103	8	.358	9
1g	.479	3	.588	5
1h	-.054	11	.201	10
1i	.515	2	.794	1
1j	.140	6.5	.455	6
1k	.746	1	.661	2
Section 2				
2a	.321	6.5	.394	10
2b	.382	5	.612	4
2c	.164	11	.548	6
2d	.249	9	.419	9
2e	.321	6.5	.443	8
2f	.697	1	.515	7
2g	.625	2.5	.843	2
2h	.176	10	.055	11
2i	.516	4	.576	5
2j	.625	2.5	.855	1
2k	.297	8	.685	3
Section 3				
3a	.780	2	.718	8
3b	.500	6	.870	3
3c	.253	10	.791	6
3d	.598	5	.876	2
3e	.141	11	.433	11
3f	.475	7	.627	10
3g	.861	1	.865	4
3h	.392	9	.835	5
3i	.687	3	.888	1
3j	.469	8	.779	7
3k	.603	4	670	9

Candidate	First Period Index	First Period Rank	Third Period Index	Third Period Rank
Section 4				
4a	.429	8	.524	5.5
4b	.786	2	.905	1
4c	.595	7	.714	2
4d	.691	4	.262	8
4e	.762	3	.286	7
4f	.810	1	.524	5.5
4g	.596	6	.596	4
4h	.667	5	.643	3
4i	.286	9	.215	9
Section 5				
5a	.750	5	.678	5.5
5b	.535	6	.428	7
5c	.893	2	.857	1.5
5d	.927	1	.821	3
5e	.857	3	.857	1.5
5f	.786	4	.678	5.5
5g	.499	7	.785	4
5h	.356	8	.356	8
Section 6				
6a	.425	9	.685	5
6b	.500	7.5	.346	10
6c	.400	10	.879	1
6d	.550	4.5	.540	8
6e	.550	4.5	.394	9
6f	.517	6	.649	6
6g	.851	1	.873	2
6h	.500	7.5	.564	7
6i	.667	3	.752	3
6j	.684	2	.734	4

example, should the candidates in a particular section all possess about the same amount of potential as suggested by the original line officer list incorporating all sixty candidates, the job of ranking those men within the section would be a much more difficult one than the job of ranking individuals in a more heterogeneous section. And too, one would expect a certain amount of variation within the sections in the mean line officer rank assigned originally. In other words, one section might be composed of men who had been ranked for the most part above average on the original line officer list, and ranking this section might be a more or less difficult task than ranking one composed of individuals ranked below average originally. Any difference in the composition of the individual sections might be expected to reduce the comparability of results obtained from one section with another.

The sections in the present sample have means and variances of original line officer list rankings that at least appear to differ, as presented in Table III. The small number of cases within each section reduces the reliability of the statistics describing the distribution, and thus, the confidence in a statistical test that would suggest that there are no differences among variances. The conservative procedure here would be to treat each section separately rather than to combine them into one distribution for

TABLE III

Means and Variances of Line Officer Ranking
of Men Within Sections

Section	Mean	Variance
1	31.36	297.19
2	35.36	346.23
3	27.63	286.27
4	36.22	240.84
5	25.63	256.48
6	27.20	320.16

consideration. An examination of Table III suggests that Sections 4 and 5 are composed of individuals more homogeneous with respect to officer potential than the other sections are, and Sections 2 and 4 appear to have individuals who are ranked below average in general by the line officers, whereas Sections 3, 5, and 6 appear to contain individuals somewhat above average on the line officer rank order.

Correlation coefficients between the judgment indices and the line officer rankings were computed for each section separately. The judgment indices were ranked within each section and compared with the ranking of the line officers within that section by the Spearman rank-difference correlation coefficient. The resulting coefficients, reflecting the relationship between the ability to judge officer potential similarly to the assessment staff and the possession of officer potential as estimated by the assessment staff, are presented in Table IV for the six sections at the first and third rating periods.

The correlation coefficients obtained again suggest that there are differences among the six sections. For Sections 4 and 5, the sections with the smallest variances, three of the four correlations are negative. The variances of the sections appear

TABLE IV

Correlations* Between Judgment Index and Line Officer Rankings
By Sections and Ranking Periods

Section	First Week	Third Week
1	+.22	+.48
2	+.75	+.52
3	+.03	+.24
4	-.17	+.03
5	-.20	-.68
6	+.08	+.14

* All correlation coefficients in this table are Spearman rank-difference.

to be highly correlated with the coefficients of correlation. The highest coefficients were obtained in Section 2 which also has the largest variance of the sections, as reported in Table III; whereas the sections with lower variance in the ranks by the line officers have lower coefficients. The correlation between variance and correlation coefficients between judgment index and adjudged officer potential is .89 for the first period and .86 for the third period. These coefficients differ significantly from zero even with an N of only six sections. T he level of ability as measured by original line officer ranks of the members of each group does not appear to be related to the size of the correlations reported in Table IV. Section 3 with a mean rank of 35.36 has a relatively large coefficient; while on the other hand, Section 4, with a mean rank of 36.22 has a relatively low coefficient. Another difference among the sections is the number of men within eacn section. Sections 4 and 5 have the smallest number of men, 8 and 9, as compared with the other sections, all of which have 11 men with the exception of Section 6 with 10 candidates. In dealing with such small numbers as these sections contain a few additional or a few fewer members will affect the reliability of any statistic to a greater extent than if there are large numbers in the groups. The reliability of the statistics computed from Sections 4 and 5

would be appreciably lower than the reliability of statistics from the other sections.

In view of these considerations it was decided to treat Sections 4 and 5 separately from the other sections. Since Sections 4 and 5 are more homogeneous with respect to officer potential as judged by the line officers and more homogeneous with respect to their ability to judge officer potential as indicated by the smaller spread in judgment indices and because the sections would yield less reliable statistics than the other sections, it was felt desirable to omit these sections from more refined analysis. The immediate effect of omitting these sections would be to restrict the conclusions to moderately heterogeneous groups.

The correlations between judgment indices and line officer rank for the remaining sections were examined for homogeneity. The test of homogeneity for correlation coefficients was applied to the four remaining correlations and the results of this test are reported in Table V. The probabilities associated with the degrees of freedom support the hypothesis that there is no difference among the correlations and that they could reasonably have been taken from a common population. The four correlations were averaged for each rating period by weighting inversely

TABLE V

Test of Homogeneity of Correlation Coefficients

Section	n - 3	First Rating Period		Third Rating Period	
		r	z'	r	z'
1	8	+.222	.224	+.483	.530
2	8	+.746	.962	+.519	.576
3	8	+.031	.030	+.236	.239
6	7	+.085	.085	+.140	.141
		Chi-square = 4.425*		Chi-square = 1.05*	

* For 3 degrees of freedom, chi-square with $P_{10} = 6.251$ (2).

according to the variance of the coefficients as suggested by Edwards (2). The average correlation for the first period was .32; and for the third period .36. The test of significance applied to these average coefficients indicated significance at the .05 level of confidence. These correlations may be interpreted as suggesting that for groups of moderately heterogeneous character, there is a relationship between the ability to evaluate others with respect to officer potential and the adjudged possession of officer potential. The value of this relationship, however, appears to be too small to be used for predictive purposes.

A more refined analysis of the data, however, suggests that judgment ability may be of greater value in eliminating poorer candidates than in evaluating average or better candidates. This point can be illustrated by considering the line officer ranks of individuals with low judgment indices. In the third period judgments, for instance, only three indices were below .20, and the line officer rankings of these men were 51, 57, and 60, the lowest possible rank being 60. In the first period there were nine indices below .33 and five of these individuals ranked below 51, out of 60, two-thirds ranked below 31, and the mean rank of all nine individuals was 41.11. Examination of the individual sections yields much the same picture. In the first section three of the four

lowest ranking judgment indices were obtained from men who had received line officer ranks of 8, 10, and 11 with the group having a lowest possible rank of 11; of the three judgments indices ranking lowest for the third period, the line officer ranks were 8, 10, and 11, respectively.

The men who exceeded judgment indices of .50 at either the first or third rating periods tended to be above average in line officer ranks. The mean rank for all such individuals was 26.23 for the total of 60 men. Sixty-four per cent of the top quarter of the line officer list came from this group with high judgment indices, and 65 per cent of the top fifth were from this group also. Taking another approach, 57 per cent of the members of this group were in the top quarter of the line officer rank and 43 per cent of these men were in the top fifth of the criterion list. It is likely then, that individuals with judgment indices above .50 will be ranked above average by line officers, with about a 50 per cent chance of being in the top quarter of the line officer list. On the other hand, individuals whose judgment index in the third period falls below .50 are more likely to receive line officer rankings that are below average, particularly if their judgment index is below .20. Half of the individuals with judgment indices below .50 were ranked lower than 41 on the final line officer list,

the mean rank of the whole group being 30.5, and only 30 per cent were in the top third of the line officer list. Those candidates with judgment indices below .20 at the third period were all ranked below 51 on the line officer list, and none appeared in the top third.

It appears, then, that an extremely low judgment index is associated with poor officer potential as judged by experienced officers; this study would suggest that a candidate will receive a low rank at the end of the screening course if his judgment index is below .20. The use of high judgment indices to identify men who will probably receive high ranks at the end of the screening course is apparently somewhat limited.

C. Improvement of Judgment from First to Third Periods

The preceding analysis suggested that the ability of an individual to judge officer potential is related to his possession of officer potential in terms of the measurement available here. This is particularly true with those men who are judged to be low in officer potential. The present section is concerned with the relationship between position on the final rank order list and the amount of improvement in judging other candidates taking place from the first rating period to the third period.

The relationship between the rankings made by each candidate at the first and third periods may be considered an indication of the reliability of the peer rankings. The average correlation from all sections was .71, the coefficients ranging from .61 to .80.

All judgment indices for both the first and third rating periods were placed in one of three categories: below .20, between .20 and .50, or above .50. The value of .20 was selected as a cutting point since data presented in the previous section indicated that values below this amount are associated with low officer potential. The .50 value was arbitrarily selected as a dividing line between average and above average judgment ability. Individuals who had indices below .20 at the first rating period are tabulated in Table VI according to the index they obtained at the third rating period. The mean line officer rank for these individuals is given separately for the three categories of judgment indices at the third rating period. The data suggest that candidates who improve, that is, have higher judgment indices in the third rating period than in the first, have higher line officer rankings. However, of the individuals who had low judgment indices in the first rating period, only those who improve considerably have line officer ranks above the mean. A similar tabulation has been made for judgment indices falling in the two higher categories on the first rating period. These tabulations are shown in Tables VII and VIII. Those candidates who had judgment indices between .20 and .50 in the first

TABLE VI

Mean Line Officer Ranks of Candidates
With Judgment Indices Below .20
At First Rating Period

Judgment Index at Third Rating Period	Number of Candidates	Mean Line Officer Rank	
		In Section	In Total Group
Below +.202	3	10.67	56.00
Between +.202 and +.500	4	7.00	36.50
Above +.500	2	5.00	28.00

TABLE VII

Mean Line Officer Ranks of Candidates
With Judgment Indices Between .20 and .50
At First Rating Period

Judgment Index at Third Rating Period	Number of Candidates	Mean Line Officer Rank	
		In Section	In Total Group
Below +.202	0	---	---
Between +.202 and +.500	5	5.91	31.80
Above +.500	12	5.00	25.75

TABLE VIII

Mean Line Officer Ranks of Candidates
With Judgment Indices Above .50
At First Rating Period

Judgment Index at Third Rating Period	Number of Candidates	Mean Line Officer Rank	
		In Section	In Total Group
Below +.202	0	---	---
Between +.202 and +.500	1	7.27	50.00
Above +.500	16	5.06	26.37

rating period, and who showed judgment indices of .50 and above on the third, have a better standing of the line officer list than those who do not improve in judgment index. This is consistent with the findings from individuals who have judgment indices below .20 in the first rating period. There was not much change in judgment index found with individuals who in the first rating period had judgment indices above .50.

The numbers involved in each of the categories are so few that the data should be taken as suggestive only. The data suggest then that in general whenever improvement in judgment takes place in the third rating period the candidate is likely to be higher in the line officer list than he would be if no improvement in the judgment index had occurred. Where no increase in judgment index from the first to the third periods is shown in the two categories where improvement is possible, the candidate is likely to be below average in line officer rank at the end of the screening course. In general, judgment indices either remained the same or increased from the first to the third rating period. Only one individual had a judgment index in the third period below his index in the first period, changing from the highest category to the middle category. The mean line officer rank of the men who stayed in the same category was 40.88, below the mean of 30.5 for the rankings of the total group of 60.

D. Summary of Findings

There are four main findings from this analysis of candidates as judgers of officer potential in their peers.

1. A low positive correlation between judgment index and line officer rank was found to be significantly different from zero at the .05 level of confidence.

2. Candidates with judgment indices above .50 on the first rating period are likely to be above average on the final rank order list; this relationship, however, is not sufficiently complete to permit exact prediction.

3. Candidates with judgment indices at the first rating period between .20 and .50 will probably be somewhat below average on the final line officer list; however, if marked increase in judgment index occurs by the third rating period, they are more likely to be above average on the line officer list.

4. Candidates with judgment indices at the first rating period below .20 are likely to be considerably below the average on the final rank order list, except in the infrequent case when third period judgment index increases to above .50.

III. SIMILARITY OF JUDGMENTS OF OFFICER POTENTIAL

In Part II of this report, the ability of a candidate to judge officer potential in his peers was evaluated by comparing his judgments with the judgments of the assessment staff. It is felt desirable to compare the ratings of peers made by one candidate with those made by other candidates in the group, that is, to consider the degree to which a candidate's judgment of his peers agrees with the judgments made by the other members of the group taken together. It is intended to determine the extent to which agreement with other candidates in evaluating peers is an indication of officer potential and to investigate those individuals whose evaluations deviate considerably from those of the group.

In this portion of the report only Sections 1, 2, 3, and 6 are used since, as indicated earlier, they contain sufficiently heterogeneous candidates to permit meaningful analysis. Sections 4 and 5 are not included in the present analysis because they give evidence of being different from the other sections.

A. Index of Similarity

An index suggested by Cronbach and Gleser (1) symbolized as $\bar{D}$, was used for the present analysis. The index reflects the

average relationship of a candidate's rating with those of other individuals in a section. Calculation of this index for each candidate would result in a distribution of $\bar{D}$'s that would fall into the same order as a distribution of rank-difference coefficients, except the reverse order would be obtained; that is, the highest $\bar{D}$ would indicate the lowest relationship. The index has the advantage of being linear in character and consequently more readily comparable. Its chief disadvantage is that an interpretation of the size of the index cannot be made in absolute terms, but can be made only in relation to the number of individuals involved; thus, the same value of the index could indicate different degrees of relationships in two situations where the size of the groups differs. In the present situation all of the groups except one were of the same size and the index is pro-rated for the odd sized group so that the reported indices are comparable for all four sections.

For the purposes of this report the index $\bar{D}$, will be referred to as the "peer-agreement index," since it reflects the similarity of an individual's ranking of members of a section to the judgment of each of the other members of the group.

B. Relationship of Agreement with Group to Adjudged Officer Potential

The index $\bar{D}$ is obtained by first taking one candidate's rankings of the others in his section and finding the squared difference between his rankings and the rankings made by another member of the section. The square root of the sum of these differences is the D for two individuals in the section. If D is computed for Candidate 1a with each of the other members of the section, and all these D's are averaged, the resultant is $\bar{D}$ for Candidate 1a. The index $\bar{D}$ was computed for each man, and within each section the individual $\bar{D}$'s were ranked. The rank order was compared by a Spearman rank-difference coefficient to the ranking that the candidate had received from the line officers at the end of the screening course. The $\bar{D}$'s, their ranks, and the line officer ranks, from all four sections for both the first and the third periods, are presented in Tables IX through XII. The rank-difference correlation coefficient reflecting the relationship between a man's agreement with his peers and his adjudged officer potential was computed for each section and each rating period and are presented in Table XIII. For the first rating period, a test (2) was applied to determine the tenability of the assumption that the four values of the correlation coefficient could be random samples

TABLE IX

Mean Peer-agreement Index and Line Officer Rank
For Members of Section 1

Candidate	Mean Index		Rank of Index		Line Officer Rank Within Section
	First Period	Third Period	First Period	Third Period	
1a	6.46	6.91	1	2	3
1b	8.83	7.21	7	3	4
1c	8.56	6.87	5	1	1
1d	9.94	9.06	10	8	2
1e	8.18	10.42	4	11	11
1f	7.93	8.41	3	5	8
1g	7.91	7.68	2	4	9
1h	10.31	8.83	11	9	10
1i	9.40	10.02	8	10	6
1j	8.82	8.67	6	7	5
1k	9.90	8.61	9	6	7

TABLE X

Mean Peer-agreement Index and Line Officer Rank
For Members of Section 2

Candidate	Mean Index		Rank of Index		Line Officer Rank Within Section
	First Period	Third Period	First Period	Third Period	
2a	8.80	9.76	6	9	5
2b	8.15	7.07	2	2	3
2c	8.58	7.59	3	4	9
2d	9.47	7.69	8	5	8
2e	8.68	9.00	5	8	10
2f	10.06	8.37	11	7	4
2g	8.60	8.21	4	6	2
2h	8.04	10.33	1	11	11
2i	9.69	10.26	10	10	1
2j	9.45	7.41	7	3	6
2k	9.48	7.02	9	1	7

TABLE XI

Mean Peer-agreement Index and Line Officer Rank
For Members of Section 3

Candidate	Mean Index		Rank of Index		Line Officer Rank Within Section
	First Period	Third Period	First Period	Third Period	
3a	8.83	5.89	7	4	6
3b	10.03	5.91	9	5	3
3c	8.06	6.00	6	6	1
3d	7.83	6.65	3	8	2
3e	10.52	9.02	10	11	11
3f	9.29	5.85	8	3	4
3g	7.91	5.03	4	1	9.5
3h	7.71	6.47	2	7	9.5
3i	7.94	5.48	5	2	7
3j	10.80	7.25	11	10	8
3k	7.46	6.97	1	9	5

TABLE XII

Mean Peer-agreement Index and Line Officer Rank
For Members of Section 6

Candidate	Mean Index		Rank of Index		Line Officer Rank Within Section
	First Period	Third Period	First Period	Third Period	
6a	7.88	9.38	9	9	9
6b	6.19	7.74	3	7	5
6c	6.24	6.40	4	2	4
6d	6.55	7.93	6	8	6
6e	6.96	9.87	7	10	8
6f	7.60	7.10	8	5	7
6g	6.37	5.30	5	1	2
6h	6.12	6.67	2	3	1
6i	8.59	7.69	10	6	10
6j	5.46	6.87	1	4	3

TABLE XIII

Correlation Between Peer-agreement Index and Line Officer Rank
By Sections and Rating Period

Section	First Period	Third Period
1	-.01	+.57
2	-.37	+.06
3	+.11	+.13
6	+.88	+.75

from a common population. The probability associated with the obtained chi-square is less than .01. The value obtained from Section 6 is the basis for permitting the rejection of the null hypothesis since the test yields a probability between .70 and .50 if Section 6 is omitted from the analysis.

Since the composition of Section 6 was similar to the other sections with respect to the mean and variance of the line officer ranking of the men in the section, some other factor or factors must be assumed to be operating that account for the difference between this section and the others. The individual judgment indices as reported in Table II appear to be considerably less variable for Section 6 than for the other sections in the first rating period. With the exception of one index obtained from Section 6, all the indices lie between .40 and .68; whereas, the indices for the other sections and other periods range from .04 to .93. In addition, the peer-agreement indices for Section 6 do not contain extreme deviates in the first rating period. The list for Section 6 contains none over 8.6; whereas, the other sections contain indices of 9 and 10. Excluding Section 6 for the first period, the average correlation of the remaining three correlations is not significantly different from zero; the correlation from Section 6 alone is not different from zero either. The evidence

TABLE XIV

Test of Homogeneity of Correlation Coefficients
For First Rating Period

Section	n – 3	r	z'
1	8	-.008	-.010
2	8	-.371	-.388
3	8	+.113	+.116
6	7	+.879	+1.376

chi-square = 14.420*

TABLE XV

Test of Homogeneity of Correlation Coefficients
For Third Rating Period

Section	n – 3	r	z'
1	8	+.574	+.655
2	8	+.065	+.065
3	8	+.132	+.131
6	7	+.746	+.962

chi-square = 4.066*

* For 3 degrees of freedom, chi-aquare with P_{01} = 11.341, and for P_{20} = 4.62 (2)

suggests, then, that the extent to which a candidate agrees with other candidates in evaluating his peers at the first rating period is not related to his adjudged officer potential.

For the correlations obtained at the end of the third period the same test (2) used for the first period ratings was applied to determine the tenability of the assumption that the correlation coefficients are random samples from the same population. The probability attached to the obtained chi-square in the present case was between .30 and .20, suggesting the tenability of the assumption. It will be noted that for the third period the judgment indices and peer-agreement indices for Section 6 do not have the same unusual characteristics as they did in the data from the first rating period. The absence of these characteristics in a situation where Section 6 did not differ from the other sections and the presence of these characteristics in the situation where differences were obtained, adds support to the speculation that the differences found earlier between Section 6 and the other sections might be related to these characteristics.

Since the correlation coefficients found at the third period can be considered random samples from a common population, the four values were averaged yielding an average coefficient of +.41,

which is significantly different from zero at better than the .02 level of confidence. It appears then that there is a significant relationship between the index of peer-agreement and the adjudged officer potential for the third rating period only. For the first period, then, no relationship between peer-agreement index and adjudged officer potential was found, but a relationship significantly greater than zero was observed for the third rating period. The hypothesis is immediately suggested that perhaps the ratings for the first period were made before group norms had been established as strongly as in the third rating period. If this were the case, it is likely that an extreme deviate at the third rating period might be one who was unable to accept the judgment of the group, a characteristic that might be related to officer potential.

C. Characteristics of Extreme Deviates

The significant correlation obtained between peer-agreement index and line officer rankings for third period ratings suggests that in general, those individuals who tend to evaluate their peers more similarly to the group evaluation have a greater tendency to rank high in adjudged officer potential, while those who evaluate their peers less similarly to the group are more likely to be ranked lower on the line officer list. The correlation coefficient itself,

however, probably does not reflect adequately the position of the extreme deviates from the group.

The frequency of occurrence of peer-agreement indices as high as or higher than 9 was sufficiently low to consider such an individual as an extreme deviate. Only ten individuals in the third rating period and twelve individuals in the first rating period had peer-agreement indices this high. In the first period five of the extreme deviates were in the bottom quarter of the line officer list, four in the top quarter, and only three fell in the middle group. In the third rating period five of the extreme deviates were in the bottom quarter, three in the top quarter, and two in the middle. One of the two indices falling in the middle group was very near to the lower limit of the top quarter.

These results would suggest for our present sample an extreme deviate is either considerably above average or considerably below average in adjudged officer potential. There is a slightly greater probability that he will be below average than above average, however.

IV. RELATIONSHIP BETWEEN PEER RATING OBTAINED AND LINE OFFICER RANK

The mean rank received by each candidate from the other members in his section might be taken as an indication of officer potential. Spearman rank-difference coefficients were computed between the ranks given to the members of each section by their peers with those ranks given by the section assessor at the end of each week. This correlation does not reflect independent opinions of the section assessors since they may use the peer ratings as information relevant to evaluating a man. Rank-difference correlations were also computed between mean peer rank received and rank on the final assessment list. The same situation with respect to non-independence exists here. The line officer assessment staff use the information from the peer ratings as well as the information from the section assessors lists in arriving at the final list. The correlation coefficients from each section for all three rating periods are presented in Table XVI.

The relationship between peer ratings and both the final and section assessors' lists increases from the first through the second to the third rating periods. The increase in relationship is not so great in the final list, however, and the average correlation from the third rating period (.76) is not essentially different from the

TABLE XVI

Rank-difference Correlations Between Rank Perceived From Peers and Position on Final and Section Lists

Final List

Section	Rating Period 1	2	3	Average
1	.55	.70	.75	.67
2	.54	.65	.75	.65
3	.71	.90	.85	.82
4	.80	.65	.65	.70
5	.98	.74	.76	.83
6	.66	.78	.78	.74
Average	.71	.74	.76	.74

Section List

Section	Rating Period 1	2	3	Average
1	.50	.65	.92	.69
2	.78	.90	.91	.86
3	.84	.79	.80	.81
4	.84	.84	.96	.88
5	.71	.88	.98	.86
6	.59	.84	.96	.80
Average	.71	.82	.92	.82

average at the first rating period (.71). It is to be expected that correlations with the line officer rank would change less than those with the section assessors' lists since three separate lists are involved for each rating period whereas the final list was issued at the end of the three weeks' period.

The correlation coefficients based on the section assessor lists tend to be higher than those based on the final lineal list. This finding is probably partly a function of the changes that had to be made in the final list to permit correlation within each section since the original list, it will be recalled, encompassed all the men in one ranking. It is possible too, that the section assessors' lists were less independent of the peer rankings than was the final list.

V. SUMMARY AND CONCI USIONS

!. Six sections of Marine Corps officer candidates prepared rankings of the other members of their sections at the end of each of three weeks. The ranking of each candidate made was compared with the ranking prepared by the assessment staff (judgment index) and compared with the rankings constructed by other members of his section (peer-agreement index).

2. For moderately homogeneous sections, a low, positive correlation was found between judgment index and standing on the final lineal list. Candidates with consistently low judgment indices are very likely to be ranked below average on the final list. The reverse is true for candidates with very high indices, but to not as great an extent. Only one of the sixty indices decreased substantially from the first to the third rating period; all the other indices either increased or remained about the same. A substantial increase in judgment index increased the likelihood of a candidate's being above average on the final list.

3. The average reliability coefficient of the peer ratings from the first to the third ratings was .71.

4. No relationship was found between the peer-agreement index and standing on the final lineal list at the end of the first week. A correlation coefficient different from zero at the .02 level of confidence was found at the end of the third week. Candidates who disagreed considerably with their peers in evaluating members of the group, were more likely to be in the upper or lower quarter of the final lineal list than in the middle of the group.

5. The mean rank a candidate received from his peer was related substantially to his rank on the final lineal list. The average correlation coefficient for all three rating periods was .74. This coefficient does not reflect independent judgments, however, since the assessment staff may have used the peer ratings in constructing the final list.

BIBLIOGRAPHY

1. Cronbach, L. J. and Gleser, Goldine, Similarity between persons and related problems of profile analysis. Technical rept. no. 2, Bureau of Research and Service, College of Education, University of Illinois, Urbana, Ill., 1952.

2. Edwards, A. L. Experimental design in psychological research. New York: Rinehart, 1950.

3. McClure, G. E., Tupes, E. C. and Dailey, J. T. Research on criteria of officer effectiveness. HRRC Research Bulletin 51-8, Lackland AFB, San Antonio, 1951.

4. Rohrer, J. H., Bagby, J. W., and Wilkins, W. L. The potential combat officer: a medicopsychological study of officer candidates for the United States Marine Corps. Technical Report. Tulane University, 1951.

5. Williams, S. B. and Leavitt, H. J., Group opinion as a predictor of military leadership, J. consult. Psychol., 1947, 11, 283-292.

6. Williams, S. B. and Leavitt, H. J. Methods of selecting Marine Corps officer candidates. In G. A. Kelley (Ed.), New Methods in Applied Psychology. College Park, Md.: University of Maryland, 1947, 96-99.

7. Wherry, R. J. and Fryer, D. H. Buddy ratings: popularity contest or leadership criteria? Personnel Psychol., 1949, 2, 147-159.