

RELIABILITY OF PEACE CORPS SELECTION BOARDS:

A STUDY OF INTERJUDGE AGREEMENT BEFORE AND AFTER BOARD DISCUSSIONS

LEWIS R. GOLDBERG¹

University of Oregon and Oregon Research Institute

The most crucial link in the Peace Corps selection process is the Advisory Selection Board, where a comprehensive pool of assessment data on each Peace Corps trainee is evaluated and discussed. In an effort to better understand this important process of collective clinical judgment, 9 Peace Corps Selection Boards were studied. Agreement between Board participants on the overall suitability of each Peace Corps trainee prior to Board discussion was compared with that reached after Board discussion. In general, the findings from 9 Selection Boards appeared remarkably similar, indicating that Board discussions tend to (a) decrease the suitability ratings for the average trainee, (b) increase the average dispersion of ratings for the group of trainees, and (c) increase quite dramatically the degree of consensus among Board participants. The unusually high consensus among participants after Board discussions attests to the rationality—though not necessarily the validity—of the Peace Corps selection process.

Of all of the personnel selection procedures now in existence, one of the most comprehensive is that utilized by the United States Peace Corps to select Peace Corps Volunteers. In an effort to be fair to each applicant and to minimize Volunteer failures overseas, the Peace Corps has instituted a 2-stage selection system. In the first stage, applicants are appraised on the basis of a detailed questionnaire, from 6 to 16 references, and a set of aptitude test scores; promising applicants are invited to participate in a 2 to 4-month period of intensive Peace Corps training.

During this training period, the second stage of selection takes place. Each Peace Corps trainee is evaluated by a host of people who have observed some aspect of his functioning; these viewpoints about each trainee stem from four major sources: (a) A Civil Service full-field background investigation gathers evaluations of the trainee's past performance in his schooling, his work, and his recreational activities. (b) Training instructors provide information on the trainee's participation and achievement in Peace Corps courses, as well

as furnishing their impressions of each trainee's potentiality as a Peace Corps Volunteer. (c) Fellow trainees furnish extensive peer evaluations, pooled judgments providing a portrait of each trainee as his peers have viewed him during the intensive training period. And, (d) staff psychologists and psychiatrists assess each trainee on the basis of their own professional tools—psychological tests and clinical interviews. Peace Corps assessment provides an attempt to understand each Peace Corps trainee through an integration of these diverse views of his behavior and effectiveness. One or more PhD psychologists, called Assessment Officers, serve at each training site. It is their responsibility to integrate the viewpoints of those who have observed the trainee's behavior during the training period and to provide a unified picture of the trainee's strengths and weaknesses for Peace Corps service overseas.

In an effort to reach the wisest possible selection decisions, all of this rich and varied material is reviewed in two staff conferences, called Advisory Selection Boards. The first, or Intermediate Selection Board, takes place approximately halfway through training; the Final Selection Board is held at the end of the training period. The Project Director, the Assessment Officers and their psychiatric consultants, all participate on these Boards,

¹ The author wishes to express his appreciation to the Selection Board members who participated in this project, and to the United States Peace Corps for providing such an important laboratory for naturalistic research. This study was supported, in part, by Grant MH-04439 from the National Institute of Mental Health, United States Public Health Service.

as do representatives from the Selection, Training, and Program-Development-and-Operations Divisions of the Peace Corps. These Boards serve much the same function as a presidential cabinet, seeking to advise the Peace Corps official—the Selection Officer—who has been entrusted with the final responsibility for making an objective and fair decision regarding the makeup of the Peace Corps group sent overseas.

Certainly the most critical component of the entire Peace Corps selection process is the Advisory Selection Board, since it is at this point that selection decisions are made. Consequently, knowledge of the reliability of judgments made during these Boards is of crucial importance to the Peace Corps. Moreover, these Boards provide an important natural laboratory for the study of decision making in small group settings, and the findings from empirical studies of these Boards may be generalizable to other small groups which are less tractable to scientific study.

The present report summarizes studies of nine Peace Corps Selection Boards, focusing on the degree of agreement between Board participants prior to, and after, Board discussion of each trainee. Consequently, this report provides evidence of the effects upon conformity in judgments of small group discussions in a natural problem-solving situation.

METHOD

Training Projects

Five different Peace Corps training projects at the Peace Corps Center of the University of Hawaii at Hilo were studied. The first project was composed of individuals training to go to Thailand, primarily as teachers of English. Pre- and postdiscussion ratings at an Intermediate Selection Board (Board A) held after 6 weeks of training, and pre-Board, prediscussion, and postdiscussion ratings at a Final Board (Board C) were studied. The second project was composed of individuals training for service in Malaysia; the group included nurses, secretaries, laboratory technicians, and three types of Rural Community Action workers (primarily individuals with agricultural backgrounds). Pre- and postdiscussion ratings at an Intermediate Board (Board B), and postdiscussion ratings at a Final Board (Board D) were studied. The third and fourth projects also were composed of individuals training to go to Thailand as English teachers. Postdiscussion ratings from a Final Board (Board E) of one project, and postdiscussion ratings from an Intermediate Board (Board F) and Final Board (Board H) of the other project, were studied. The fifth project was also bound

for Thailand, as village health and sanitation workers; postdiscussion ratings from an Intermediate Board (Board G) and a Final Board (Board I) were studied.

Board Participants

The composition of Peace Corps Selection Boards differs slightly, depending at least in part upon (a) the Selection Officer's choice of persons to whom he wants to turn for advice, and (b) the availability of these individuals to meet together at a common time at the training site. The Selection Officer, who ultimately has the responsibility for Peace Corps selection decisions, is always present at these Boards. For all nine of the Boards under study, the same Selection Officer (the author) was involved.

Five different Assessment Officers participated on one or more of these Boards. In addition, three Assessment Associates, graduate students in psychology taking time out from their studies to work full time for the assessment staff, participated on one or more of the Boards, as did two psychiatrists. The Project Director—the administrator of the training project—participated on all nine Boards, and two of his assistants were also involved on some of the Boards. Three Coordinators—members of the instructional staff having responsibility for supervising the technical training in the project—also participated on one or more Boards. From the Peace Corps Washington Staff, two Training Officers participated, each on a different Board. In addition, four Desk Officers from the Program-Development-and-Operations Division of the Peace Corps each participated on a Board. Three members of the Peace Corps overseas staff were also involved on one or more of the Boards.

Rating Scale

All Board participants rated each trainee on an 11-point rating scale, where "0" indicated that the trainee should be separated from the project at the time of the Board and "10" indicated an "ideal Peace Corps Volunteer." Participants were free to use the intermediate points on the scale as they wished.

Procedure

All Board participants were presented with a set of mimeographed material summarizing the data collected during training. Approximately two typewritten pages of information, summarizing the trainee's grades, staff evaluations, peer ratings on 6–12 traits, psychiatric notes, and the assessment staff's summary evaluation, were presented for each trainee. In addition, many of the participants had some information not shared by others; for example, only the Selection Officer had read the background reports, while members of the administrative and assessment staff had had personal contact with most of the trainees. While the MMPI and a few less-structured assessment instruments (e.g., a sentence-completion form) were administered to all trainees in each of the five projects, these data were seen only by the assessment staff. The only test scores available to all Board participants were scores

TABLE 1
 AGREEMENT AMONG SEVEN SELECTION BOARD PARTICIPANTS BEFORE AND AFTER BOARD DISCUSSION
 (BOARD A; 57 TRAINEES)

	Selection officer	Assessment officer-1	Assessment officer-2	Psychiatrist-1	Project director	Coordinator-1	PDO officer-1	Prediscussion M	Prediscussion σ	Prediscussion \bar{r}
Selection officer	.97	.78	.69	.65	.78	.59	.63	5.0	2.6	.69
Assessment officer-1	.91	.89	.83	.76	.78	.57	.59	5.3	1.9	.72
Assessment officer-2	.92	.89	.83	.69	.71	.62	.62	6.0	1.6	.69
Psychiatrist-1	.80	.83	.72	.89	.75	.70	.54	5.2	2.4	.68
Project director	.86	.86	.84	.87	.87	.65	.63	5.2	1.6	.72
Coordinator-1	.91	.90	.91	.81	.86	.76	.39	6.3	1.5	.59
PDO officer-1	.74	.68	.72	.69	.76	.71	.86	4.8	1.4	.57
Postdiscussion M	4.8	5.0	5.7	5.0	5.0	5.8	4.6			
Postdiscussion σ	2.7	2.5	2.4	2.7	2.5	2.1	2.0			
Postdiscussion \bar{r}	.86	.84	.83	.79	.84	.85	.72			

Note.—All 57 trainees were rated on a 0–10 scale. Correlations among ratings *prior* to Board discussion are presented above the diagonal; correlations among ratings *after* the Board discussion are presented below the diagonal. Diagonal entries (circled) present the pre- versus postdiscussion correlation for each participant.

on a modern language aptitude test and scores on a test of general verbal ability.

For Boards at which prediscussion ratings were collected (Boards A and B), Board participants were requested to read the assessment material of the first trainee and on the basis of this information, plus any other idiosyncratic information they might possess, to rate the trainee's overall suitability for service as a Peace Corps Volunteer. After all ratings were completed for the first trainee, the Board discussed the trainee and any idiosyncratic information was shared and evaluated. When the discussion was completed, but before turning to the next trainee, postdiscussion ratings were made. The same procedure was followed for each trainee, in turn. The ratings were made individually, and the raters did not compare their ratings with each other. The procedures for Board C were the same as for Boards A and B, with one exception: for Board C, all participants (except the Selection Officer) rated each trainee *prior* to the actual start of the Board, before any of the mimeographed assessment material was distributed. These pre-Board ratings, based upon highly variable amounts of personal observation of the trainees during the training period, were then followed by the same prediscussion and postdiscussion rating procedures used with Boards A and B. The procedures for the other six Boards differed only in

that no prediscussion (and no pre-Board) ratings were made.

RESULTS

Table 1 summarizes the findings for Board A. In Table 1 the prediscussion correlations among ratings are presented above the diagonal, while the corresponding postdiscussion correlations are presented below the diagonal. Note that for every pair of Board participants, postdiscussion correlations were higher than prediscussion ones, suggesting that some appreciable consensus in ratings occurred as a function of Board discussions. The mean prediscussion correlation was .66; the mean postdiscussion correlation was .82. Moreover, all of the standard deviations of the ratings were slightly higher after discussion ($\bar{\sigma} = 2.4$) than before discussion ($\bar{\sigma} = 1.9$), indicating that discussion served to increase judgmental differentiation among the trainees being rated.

Of some interest is the marked stability of prediscussion versus postdiscussion ratings

TABLE 2
 AGREEMENT AMONG SEVEN SELECTION BOARD PARTICIPANTS BEFORE AND AFTER BOARD DISCUSSION
 (BOARD B; 88 TRAINEES)

	Selection officer	Assessment officer-3	Assessment associate-1	Psychiatrist-1	Project director	Coordinator-2	Training officer-1	Prediscussion M	Prediscussion σ	Prediscussion \bar{r}
Selection officer	(.96)	.82	.83	.79	.80	.67	.78	5.1	2.5	.78
Assessment officer-3	.91	(.89)	.82	.78	.82	.70	.83	6.2	1.7	.80
Assessment associate-1	.92	.88	(.93)	.80	.80	.70	.82	5.3	2.3	.80
Psychiatrist-1	.85	.83	.86	(.90)	.73	.70	.76	5.0	2.2	.76
Project director	.86	.88	.86	.82	(.87)	.73	.71	5.5	1.7	.76
Coordinator-2	.80	.84	.80	.80	.83	(.87)	.61	5.5	1.5	.68
Training officer-1	.88	.87	.85	.83	.81	.71	(.86)	5.7	2.3	.75
Postdiscussioin M	4.9	5.8	5.1	5.2	5.5	5.6	5.5			
Postdiscussioin σ	2.6	2.4	2.7	2.5	2.2	1.9	3.1			
Postdiscussioin \bar{r}	.87	.87	.86	.83	.84	.80	.82			

Note.—All 88 trainees were rated on a 0-10 scale. Correlations among ratings *prior* to Board discussion are presented above the diagonal; correlations among ratings *after* the Board discussion are presented below the diagonal. Diagonal entries (circled) present the pre- versus postdiscussion correlations for each participant.

for individual Board participants. While the average of these correlations was .87, the Selection Officer seemed unusually resistant to change ($r = .97$). One explanation for both the apparent convergence onto the Selection Officer's ratings and his own seeming stubbornness may stem from the fact that the Selection Officer had access to the background reports, the tenor of which he conveyed to the Board during the discussion period.

A common observation among Selection Board participants is that a great deal of the discussion of each trainee focuses on undesirable or derogatory aspects of his behavior. An indirect confirmation of this impression can be seen by comparing the prediscussion versus postdiscussion *mean* ratings. While none of these changes was great, there was a consistent tendency for the prediscussion mean ratings ($M = 5.4$) to be higher than the corresponding postdiscussion means ($M = 5.1$),

suggesting a slight devaluation of the suitability of the average trainee after Board discussion.

Table 2 summarizes the corresponding findings for Board B. Virtually all of the findings from the analyses of Board A were replicated in Board B, though the correlation coefficients tended to be somewhat higher. The mean prediscussion correlation was .76, while the mean postdiscussion correlation rose to .84. Again, the average postdiscussion standard deviation ($\bar{\sigma} = 2.5$) was higher than the corresponding prediscussion one ($\bar{\sigma} = 2.0$). Differences in mean ratings were slight, though again the average postdiscussion mean (5.5) was higher than the average prediscussion mean (5.4). Intrajudge correlations, pre- versus postdiscussion, were again remarkably high ($\bar{r} = .90$), with the Selection Officer again being the most recalcitrant to change ($r = .96$).

In order to check whether the consensus

TABLE 3
 AGREEMENT AMONG EIGHT BOARD PARTICIPANTS PRIOR TO RECEIVING FINAL BOARD INFORMATION
 (BOARD C; 46 TRAINEES)

	Assessment officer-1	Assessment officer-2	Assessment associate-1	Psychiatrist-1	Project director	Assistant director-1	Coordinator-1	Overseas representative
Assessment officer-2	.54							
Assessment associate-1	.13	.67						
Psychiatrist-1	.20	.67	.43					
Project director	.22	.38	.16	.17				
Assistant director-2	.48	.54	.43	.36	.35			
Coordinator-1	.51	.70	.58	.48	.45	.42		
Overseas representative	.16	.60	.41	.34	.39	.27	.50	
<i>M</i>	5.1	4.8	5.7	5.3	6.6	5.1	6.9	5.1
σ	1.6	2.8	1.9	2.1	1.3	2.0	1.4	1.7
Average <i>r</i>	.32	.59	.40	.38	.30	.41	.52	.38
Correlation with prediscussion rating	.84	.98	.88	.84	.38	.78	.78	.56
Correlation with postdiscussion rating	.64	.88	.77	.63	.48	.64	.73	.51

among Board participants would be as great when the amount of information available for each trainee was increased, five Final Selection Boards were studied. Tables 3 and 4 summarize the findings from the first of these, Board C. Table 3 presents the correlations among eight Board participants prior to the start of the Board, before Board materials were made available. Consequently the agreement coefficients in Table 3 represent the amount of consensus among individuals after highly variable amounts of personal interaction with the trainees. Since many of the participants had had no personal contact with some of the trainees (e.g., the psychiatrist had only interviewed 28 of the 46 trainees), the correlations presented in Table 3 were computed on all of the trainees rated by both members of each pair of raters. The number

of cases used to compute each correlation varied from 12 to 46; in general, most of the correlations in Table 3 were based on 30 to 40 cases.

Comparing Table 3 with Table 4, one can see an orderly progression in judgmental convergence as more information became available. The average agreement correlation increased from .41 (pre-Board) to .68 (prediscussion) and then again to .83 (postdiscussion), illustrating rather dramatically the impact of the Board materials and Board discussion upon these suitability judgments. As before, the mean ratings tended to decrease, from 5.6 (pre-Board) to 5.1 (prediscussion) to 4.9 (postdiscussion). Judgmental differentiation among individual trainees again showed an orderly increase: from a σ of 1.8 (pre-Board) to 2.2 (prediscussion) to 2.8 (post-

TABLE 4
 AGREEMENT AMONG NINE FINAL SELECTION BOARD PARTICIPANTS BEFORE AND AFTER BOARD
 DISCUSSION (BOARD C; 46 TRAINEES)

	Selection officer	Assessment officer-1	Assessment officer-2	Assessment associate-1	Psychiatrist-1	Project director	Assistant director-1	Coordinator-1	Overseas representative	Prediscussion \bar{M}	Prediscussion σ	Prediscussion \bar{r}
Selection officer	.97	.85	.73	.76	.77	.60	.60	.74	.61	4.3	3.0	.71
Assessment officer-1	.91	.91	.71	.75	.73	.68	.57	.74	.58	4.9	2.3	.70
Assessment officer-2	.91	.90	.89	.77	.72	.68	.68	.74	.80	4.9	2.8	.73
Assessment associate-1	.86	.92	.89	.91	.67	.75	.67	.73	.68	4.8	2.6	.72
Psychiatrist-1	.83	.88	.87	.87	.88	.58	.70	.57	.58	5.0	2.2	.66
Project director	.69	.77	.77	.71	.79	.83	.58	.67	.61	5.4	2.0	.64
Assistant director-1	.76	.78	.81	.81	.80	.72	.86	.58	.66	4.8	2.0	.63
Coordinator-1	.83	.86	.86	.87	.83	.78	.80	.87	.69	6.7	1.7	.68
Overseas representative	.82	.86	.86	.86	.87	.76	.88	.81	.74	5.2	1.6	.65
Postdiscussion \bar{M}	4.2	4.5	5.1	4.5	4.9	5.1	4.5	6.3	4.8			
Postdiscussion σ	3.1	3.3	3.3	2.7	2.4	2.9	2.5	2.3	2.3			
Postdiscussion \bar{r}	.83	.86	.86	.85	.84	.75	.80	.83	.84			

See Note at bottom of Table 2.

discussion). And, as before, the pre- versus postdiscussion correlations for individual participants were high ($\bar{r} = .87$), with the Selection Officer again showing the least change ($r = .97$).

The correlations among Board participants reported in Tables 1, 2, and 4 appear unusually high compared to those previously reported in the literature on clinical judgment. Since it is possible that the prediscussion ratings tended to influence the participants and thereby confound the postdiscussion ratings, it is of theoretical interest to compare these ratings with some made after Board discussions but *without* prediscussion ratings. Consequently, six other Boards—two Intermediate (Boards F and G) and four Final (Boards D, E, H, and I)—where prediscussion ratings were not carried out, were studied. While there were some slight differences between the findings from these six Boards, on the average

their results replicated rather precisely the postdiscussion findings from the other three Boards. Table 5 presents a tabular summary of the data from all nine Boards, thus permitting an easy comparison. In general, the findings look remarkably consistent across the nine Boards.

As a check to see whether there were any differences in the degree of agreement between pairs of Board participants holding different roles (e.g., to ascertain whether psychiatrists tended to agree more with Assessment Officers, for example, than they did with the Project Director), Board participants were classified into eight role-types (Selection Officer, Assessment Officers, Assessment Associates, Psychiatrists, Project Director, Assistant Directors and/or Coordinators, PDO Officers, and Overseas Representatives) and the agreement correlations among all individuals falling into each of the resulting 28 role pairs were averaged

TABLE 5
SUMMARY OF THE ANALYSES FOR NINE SELECTION BOARDS

	Board A	Board B	Board C	Board D	Board E	Board F	Board G	Board H	Board I	Average
Type of Board	Inter- mediate	Inter- mediate	Final	Final	Final	Inter- mediate	Inter- mediate	Final	Final	
No. participants	7	7	9	8	9	7	6	8	7	
No. trainees	57	88	46	76	82	81	32	73	29	
Pre-Board <i>M</i>	—	—	5.6	—	—	—	—	—	—	5.6
Prediscussion <i>M</i>	5.4	5.5	5.1	—	—	—	—	—	—	5.3
Postdiscussion <i>M</i>	5.1	5.4	4.9	5.2	4.4	5.3	4.6	5.3	5.6	5.1
Pre-Board $\bar{\sigma}$	—	—	1.8	—	—	—	—	—	—	1.8
Prediscussion $\bar{\sigma}$	1.9	2.0	2.2	—	—	—	—	—	—	2.0
Postdiscussion $\bar{\sigma}$	2.4	2.5	2.8	2.7	2.8	2.2	2.4	2.4	2.3	2.5
Pre-Board r	—	—	.41	—	—	—	—	—	—	.41
Prediscussion r	.66	.76	.68	—	—	—	—	—	—	.70
Postdiscussion r	.82	.84	.83	.87	.76	.78	.81	.79	.75	.81
Pre- versus Post- r	.87	.90	.87	—	—	—	—	—	—	.88

across all nine Boards. In general, role differences in degree of agreement were greater for prediscussion ratings than for postdiscussion ratings. The highest postdiscussion average agreement correlations occurred between the Selection Officer and the Assessment Officers ($\bar{r} = .87$), followed closely by that between the Selection Officer and the Psychiatrists ($\bar{r} = .86$); this finding may have quite limited generality, however, since only one Selection Officer was studied. The lowest

postdiscussion average correlations were found between PDO Officers and Coordinators ($\bar{r} = .73$), and between the Assessment Associates and the Project Director ($\bar{r} = .75$).

While the preceding analyses have all focused upon ratings made during a relatively brief interval of time, of equal interest are the correlations between suitability ratings made by the same Board participants (a) at an Intermediate Board and (b) 6 weeks later at a Final Board. Since the present study includes both Intermediate and Final Board ratings from each of four projects, it was possible to compute the correlations between the two sets of ratings for those individuals who participated on both Boards. While making their Final Board ratings, Board participants did not have access to their Intermediate Board ratings.

Table 6 summarizes these findings from the three largest projects, for the subset of trainees who remained in each project until the Final Board. Included in Table 6 are the number of trainees rated at each of the Final Boards, as well as the number of Board participants who were involved in both Boards. The mean ratings and dispersions for each of these Board participants were computed across the subset of trainees they rated at both Boards; the

TABLE 6
INTERMEDIATE BOARD VERSUS FINAL BOARD RATINGS

Variable	Board A versus Board C	Board B versus Board D	Board F versus Board H
No. trainees at Final Board	46	76	73
No. overlapping Board participants	6	6	5
<i>M</i> (Intermediate Board)	5.8	5.7	5.4
<i>M</i> (Final Board)	5.0	5.1	5.4
$\bar{\sigma}$ (Intermediate Board)	2.1	2.0	2.3
$\bar{\sigma}$ (Final Board)	2.9	2.8	2.5
<i>r</i> : range	.56-.78	.55-.77	.49-.77
r	.69	.65	.65

averages of these means and standard deviations are presented in Table 6, for both Intermediate and Final Board ratings. Finally, the correlations between Intermediate and Final Board ratings were computed for each of the Board participants across the subset of trainees common to both Boards; the ranges of these correlations are presented in Table 6, along with the mean correlations.

The correlations between ratings made at Intermediate and Final Boards averaged .69, .65, and .65 for the three projects, indicating considerable—though far from perfect—judgmental stability over time. The correlations were of approximately the same magnitude as those found among participants prior to Board discussion (see Table 5).

In general, the mean ratings of the same trainees by the same Board participants were lower at the Final Board than at the Intermediate Board, while the dispersions of the ratings were larger at the Final than at the Intermediate Boards. Since some of the lowest rated trainees had been separated from the projects at the Intermediate Boards, the mean Intermediate Board ratings were higher for the subgroup which remained in the project (Table 6) than for the original groups (Table 5); conversely, the dispersions of the ratings were smaller for the Final Board subgroups than for the larger groups. At the Final Boards, however, the values of the means and standard deviations for the smaller groups approached those from the larger groups, indicating that Board participants tended to evaluate trainees relatively (e.g., "on a curve"), rather than absolutely. Since the evaluation of a particular trainee appears to be highly dependent on his relative status in his group, the validity of these Selection Board ratings could be severely attenuated by any interproject differences in the "quality" of the average Peace Corps trainee.

DISCUSSION

By far the most remarkable finding from the present study was the substantial degree of interjudge agreement among Peace Corps Selection Board participants, even prior to any Board discussion. Moreover, the increased consensus as a result of Board discussion certainly lends credence to the rationality (Goldberg,

1963) of this stage of the Peace Corps selection process. In addition, since most of the Assessment Officers on these Boards could have served as Selection Officers on other projects, the present study supplies some *indirect* evidence that the evaluations from one Selection Officer may show a substantial degree of agreement with those of others. This does not imply, however, that individual Selection Officers may not differ appreciably in their subjective cutoff points between persons selected as Volunteers and those not sent overseas. Informal discussions among Selection Officers have convinced the author that the degree of agreement on the overall ranking of a set of trainees by different Selection Officers would probably be quite great, but that Selection Officers would differ significantly in the percentage of the group they would select as Volunteers.

The degree of interjudge agreement among members of decision-making teams in other settings has never been adequately explored and, consequently, any comparison of the post-discussion correlations among Peace Corps Selection Board participants with those from other groups must await similar studies in different contexts. However, the prediscussion correlations from this study can be compared with the numerous studies of consensus in clinical judgments more generally. Such studies have indicated a vast range of reliability coefficients, depending on the nature of the judgmental task. For example, in an early inferential reliability study, Bendig (1955) asked 40 graduate students in psychology to rate each of 10 abstracted clinical case histories on a 7-point scale of global adjustment level. Bendig reported average interjudge reliability coefficients around .84. On the other hand, Howard (1962) had seven clinical psychologists rank order 10 needs for each of 10 patients on the basis of Rorschach, TAT, and sentence-completion test protocols. For this task, Howard found interjudge agreement correlations, for the same projective test, to average only .19. Clearly, the findings from the present study resemble those of Bendig more than those of Howard.

One possible explanation for the rather high consensus among judgments in the present study may stem from the fact that the overall

suitability ratings made by participants in Peace Corps Selection Boards are, by and large, mostly *evaluative* in nature (as were, of course, the ratings in Bendig's study). Were more specific predictions demanded of the Board participants, it would not be surprising if the resulting interjudge correlations would decrease appreciably.

Finally, the findings from this study should in no way be construed as reflecting on the validity of these Board ratings, since even the most reliable of clinical judgments may be badly misaligned with reality. Only through further research on the *validity* of Peace Corps selection procedures—including studies in which all trainees are sent overseas—can this

latter (and more important) question be answered definitively.

REFERENCES

- BENDIG, A. W. Rater experience and case history judgments of adjustment. *Journal of Clinical Psychology*, 1955, **11**, 127-132.
- GOLDBERG, L. R. Peace Corps selection as an assessment strategy. *Proceedings of the Peace Corps—NIMH conference: "The Peace Corps and the behavioral sciences."* Washington, D. C.: Peace Corps, 1963.
- HOWARD, K. I. The convergent and discriminant validation of ipsative ratings from three projective instruments. *Journal of Clinical Psychology*, 1962, **18**, 183-188.

(Received August 2, 1965)