

ASSESSMENT OF OSS PERSONNEL

By HENRY A. MURRAY

MEDICAL CORPS, A.U.S.

AND

DONALD W. MACKINNON

OFFICE OF STRATEGIC SERVICES

THE TASK of the assessment board of the Office of Strategic Services was that of judging the suitability of each candidate for a proposed assignment overseas, and of placing him in one of five categories: *not recommended; doubtful; recommended with qualifications; recommended; highly recommended*. During the early months of the program the candidate's personality was assessed solely in relation to the mission selected for him by the recruiting and branch officers. Later, after it was discovered that many men, on arrival in the overseas theater of operations, were assigned jobs for which they had not been assessed, an additional practice was instituted: that of assessing a candidate's fitness in relation to nine large job categories as well as in relation to the particular job proposed for him.

As a servicing agency the chief aim of the assessment board was to discriminate unfitness, and so, in this respect, its role could be compared to that of the conventional screening unit, were it not for two significant differences: first, the OSS board had to appraise the relative usefulness of men and women who fell, for the most part, in the middle and upper ranges of the distribution curve of general effectiveness or of one or another special ability, people who had already been selected because of the skill they had manifested in some field of activity; and, second, these men and wom-

en had to be assessed in respect to a number of additional personality qualifications — social relations, leadership, discretion (disposition and ability to withhold confidential information), and so forth. The usual screening devices, based on elementalistic principles, though efficient in distinguishing people who, because of one or another marked defect or handicap, are incapable of functioning effectively, are not so successful in discriminating degrees of effectiveness among those whose test scores fall above the usual level of acceptability. Furthermore, none of the conventional screening devices are designed to test social qualifications—good will, tact, teamwork, freedom from disturbing prejudices, freedom from annoying traits, and so forth.

Also, besides eliminating the candidates who did not measure up to the relatively high OSS standard of overall acceptability, the assessment board was expected to predict the performance level of each candidate in fulfilling a designated role in a designated area of operations (e.g., branch chief in Cairo, leader of native resistance groups in Burma, etc.) and, in addition, to write a personality sketch of the candidate, which would assist his superior officer, or supervisor, in dealing with him, or in deciding to which specific place, specific group, and specific function he should be assigned.

To accomplish all this it was consid-

ered necessary to set up a program of assessment according to organismic (gestalt) principles, that is to say, to design a variety of tasks-in-situations which would test a man's effectiveness in performing functions of the same type and at the same integrative level, and under somewhat similar conditions, as those he would be required to perform in the field. For OSS personnel this meant, among other things, that a candidate would have to be confronted by a succession of *varied novel* assignments—requiring quick planning, coordination of muscles with or without instruments and/or the coordination of ideas and words—many of which assignments would have to be executed in collaboration with other people. Make-believe, yet to a degree, realistic test situations conform to organismic principles in so far as they call for functional operations, or *proceedings*, of personality at a relatively high level of differentiation and integration, the level that must be sustained if one is to act appropriately and efficiently in everyday life, especially in a policy-making, administrative or executive capacity.

The assessment board also complied with organismic principles in so far as they attempted in each case to synthesize significant observations of behavior into partial formulations of the personality of the candidate, and from these formulations to deduce the behavior that could be expected of him in this or that situation in the theater. In other words, the task of assessment, as an intellectual proceeding, was a matter of diagnosis (inferring general tendencies and traits and their interrelations from a number of specific signs), and prognosis (on the basis of what is known about these tendencies, traits and relations, to predict future patterns of behavior) and finally, with the latter in mind, to decide on the suitability of

the candidate. This differs from the elementalistic mode of appraisal which consists of a series of objective measurements of relatively simple processes or isolated traits, which leads to a final decision based on a list of scores weighted in connection with known correlations between scores and levels of performance (as determined in previous studies of similar groups). The elementalistic method does not call for a diagnosis of the personality.

The bulk of OSS assessment in the United States was done at Station S which was located on a hundred-acre farm, fifteen miles from Washington, and at Station W in Washington, D.C. The former had a three-and-a-half day program, the latter a one-day program of procedures. This paper is limited to an account of the conditions that prevailed and the techniques that were employed at Station S.

There a senior staff of about 7 psychologists, psychiatrists, and sociologists, and a junior staff of about 8 graduate students of psychology lived intimately with groups of approximately 18 candidates, who, in the course of their three and a half days in the Area, were entertained and tested and assessed in a variety of ways.

Regardless of military rank or civilian status, all candidates arrived at S in army fatigues and during the period of their stay hid their identities from each other by means of fictitious names and stories of their lives.

The final evaluation of each candidate and the judgment of his suitability for his proposed assignment came from a number of sources and procedures:

1. Casual observations. These were made constantly, since staff and candidates lived together in one house, ate in the same hall, relaxed in a common livingroom, and played games together

both in- and out-of-doors.

2. Standardized tests of abstract intelligence, of mechanical comprehension, etc. These were paper-and pencil tests of both group and individual type.

3. Projective tests, such as Sentence Completion, Thematic Apperception Test, Rapid Projection Test, the Rorschach, and so forth, used primarily to uncover repressed tendencies which might affect a candidate's motivation or emotional stability.

4. Personal history interviews with each candidate lasting at least an hour and a half, in preparation for which the interviewer examined a Personal History Form filled out by the candidate, his standard test scores, and his performance on the projective tests, and in the course of which the interviewer sought insight into the structure and dynamics of the candidate's personality and his motivational, emotional, and intellectual fitness for his assignment.

5. Group situational tests in which 4 to 7 candidates were faced with the task of solving a problem, either in-doors or outdoors, in some cases with a group leader assigned and in other situations with no leader designated. Group situations of this sort were first introduced by the British W.O.S.B.'s (War Officer Selection Boards) in their testing of officer candidates.

In Leaderless Group Situations the development of group structure could be observed as well as the contribution of each man to it, his practical intelligence, his leadership, his energy and initiative, his social relations, and in outdoor situations, his physical ability. Typical leaderless group situations at S were:

a) Wall Test. A group of 6 men were brought to one of two parallel walls, 8 feet apart and 10 feet high, and told that their task was to get themselves and a "king-size bazooka" (actually a heavy 8-foot log) over the two

walls and down on the far side without getting down between the walls.

b) Discussion Test. A group of six men, sitting around a table, were asked to discuss for half an hour the question, "What are the major postwar problems which the United States will face after the war?" and at the end of that time, to choose one of their group to present their conclusions to the staff.

In Assigned Leadership Problems, each member of the group was, in turn, assigned to lead his group in solving some problem, for example, crossing a mined road, searching an agent's room in enemy-occupied territory, and so forth, situational tests in which fact and fancy were fused.

6. Individual Situational Tests in which a candidate met a situation alone, such, for example, as:

a) Construction Test. The candidate was shown a glorified tinker-toy with which he was instructed to build a 5-foot cube with 7-foot diagonals in ten minutes with the aid of 2 helpers whom he was to direct. The "helpers" were members of the junior staff who soon turned the situation into a test of the candidate's frustration-tolerance by becoming increasingly lazy, recalcitrant, and insulting. No candidate ever completed this task, and there were those who became either markedly upset or enraged by the humiliations they suffered.

b) Stress Interview. The candidate was told to assume that he had just been caught going through a secret file in a government building in Washington, and, after ten minutes in which to think up an innocent and plausible explanation for his presence there, he reported to a basement room where, under the glare of a spotlight, he was given a grilling cross-examination on the details of his story in an attempt to confuse and disquiet him as much as possible.

7. An obstacle course to test the candidate's physical ability, especially his daring and agility.

8. Tests of observing and reporting, for example, the terrain at S, personal effects left in a room, a map, contents of a brief case, etc.

9. Tests of propaganda skills, for example, a script for a spot radio broadcast and a pamphlet to disturb Japanese workers in Manchuria.

10. Psychodrama. Situations tailor-made for each candidate to reveal his ability to cope with difficult social situations *vis-a-vis* another candidate.

11. Debate. Toward the end of the program a convivial evening party was held with hard liquor for both candidates and staff with a more or less informal discussion or debate by the candidates on a timely topic, such as, "What are we fighting for?" "What should the Allies do with Germany after the war?" etc. The party which lasted usually long past midnight often revealed aspects of the candidates' personalities not seen in soberer states.

12. Sociometric questionnaire given candidates on the last day which revealed their opinions of the colleagues, those whom they would choose as their leaders, those whom they would accept socially and those whom they would reject, those who were most dogmatic, most inconspicuous, most irritating, etc.

13. Judgment of others. As a test of their ability to size up others each candidate was asked to write a personality sketch for each of the five men in the group whom he had come to know best during his stay at S.

Material from all these sources, interpreted in the light of the picture of the man obtained in the personal history interview, and, in turn, correcting that picture, was utilized in writing the Evaluation Note on the candidate which constituted the main part of the S Fi-

nal Report. In this note of six to seven hundred words an attempt was made to convey in nontechnical terms the main features of the character structure of the candidate.

A second part of the Final Report consisted of ratings on a six-point scale (very inferior; inferior; low average; high average; superior; very superior) of ten variables of personality, the first seven of which were considered important for all candidates, the last three of which were necessary for certain types of assignment.

These variables were:

1. *Motivation for Assignment*: the degree to which a man felt impelled to contribute to the war effort; his interest in the proposed assignment.
2. *Energy and Zest*: average level of purposive activity and initiative.
3. *Practical Intelligence*: skill in dealing with things, with people and with ideas.
4. *Emotional Stability*: governance of emotions under stress; absence of anxiety, anger and depression; absence of neurotic symptoms.
5. *Social Relations*: ability to get along with all kinds of people; cooperativeness; absence of annoying traits.
6. *Leadership Ability*: ability to organize groups and evoke cooperation and respect.
7. *Security*: ability to keep secrets and maintain cover.
8. *Physical Ability*: agility, daring, ruggedness.
9. *Observation and Reporting*: ability to observe and remember significant facts; to evaluate information; to report accurately and concisely.
10. *Propaganda Skills*: ability to devise and execute subversive measures for disintegrating the morale of the enemy.

In order the more adequately to observe candidates and to rate them on the variables of personality relevant for their assignments each class of 18 was divided into 3 subgroups of 6 men, to which a team of 2 senior and 1 junior staff members was assigned. It was the responsibility of this staff team to interview their 6 candidates and to rate them in various situations and tests, and to prepare the Final Report on each of them. This provided that the ratings of all personality variables of each candidate rated in a situation was the result of at least 3 independent ratings combined with final ratings in substaff conferences in which the performance of each candidate and its meaning were discussed at some length.

Since there were, in all, 35 procedures, and since each procedure provided an opportunity to rate 1 to 4 variables, there were anywhere from 2 to 10 scores for every candidate on each of the 10 variables by the end of the three and a half days of assessment. These ratings for each candidate were represented by colored pins placed on a huge scoreboard at one end of the staff room. For every rating above average a red pin was inserted, and for every rating below average a blue one, thus yielding an easily grasped picture of each man's performance in all procedures at S at the time his case was discussed by the entire staff at its final meeting on the afternoon and evening of the last day of a class.

At this staff meeting it was the custom to have the first part of the Evaluation Note on a candidate read by the senior staff member who had interviewed him and who was responsible for his overall evaluation, and to have the second part, which described the candidate's behavior in the situations at S, read by the junior staff member who had been responsible for this part of the

note. Other staff members contributed their impressions of the candidate in question, and proposed changes of or additions to the note. When the final form of the Evaluation Note had been agreed to by the staff as a whole, the final ratings of the variables for this candidate were made and a decision reached as to his suitability for his proposed assignment. The Final Reports made up of these three parts—the job fitness rating, the ratings of the 10 personality variables, and the Evaluation Note—were in this way prepared for each candidate, typed, and already delivered in Washington before the candidates reported there the next morning.

About 5,500 men and women went through the Assessment Centers in this country. Of these, about 20 per cent were not recommended for overseas duty, and another 20 per cent were not sent overseas for other reasons. Of those who did go overseas, about 1,200 were followed up in the different theaters — ETO, METO, IBTO, and CTO—where they were appraised by those who had supervised their work, as well as by 3 or 4 of their associates.

The analysis of this appraisal has not been completed. At the present time a report can be made only on the first 300 cases. On this basis, 6 out of every 100 men passed by the assessment staff proved to be unsatisfactory in the theater for one reason or another. Some of these unsatisfactory persons had been given assignments in the field other than those for which they had been assessed. Some others had failed to measure up to the high OSS standard because they lacked the necessary degree of some very specialized skill, the determination of which did not fall within the province of the assessment program. Only one of the 300 cases failed because of a neuropsychiatric condition.