



HOW ASSESSMENT CENTERS WERE STARTED IN THE UNITED STATES

THE OSS ASSESSMENT PROGRAM

A MONOGRAPH BY
DONALD W. MACKINNON

TABLE OF CONTENTS

HOW ASSESSMENT CENTERS WERE STARTED IN THE UNITED STATES

1. The OSS Assessment Program 2

REFERENCES16

ABOUT THE AUTHOR17

HOW ASSESSMENT CENTERS WERE STARTED IN THE UNITED STATES

THE OSS ASSESSMENT PROGRAM

Five months before the attack on Pearl Harbor, President Roosevelt created yet another of his alphabetical bureaucracies. This time it was the Office of the Coordinator of Information (COI), and at its head he placed General William Joseph Donovan, World War I hero and holder of the United States' three highest military decorations. Known since his youth as "Wild Bill," Donovan was an Irish Catholic, a Hoover Republican, a millionaire Wall Street lawyer, but above all he was a man of enormous energy and imagination.

His task was to lead the "New Deal's" excursion into espionage, sabotage, "black" propaganda, guerrilla warfare, and other "un-American subversive practices." (Smith, 1972, p. 1) Donovan was an understandable choice to head the COI, since he had convinced Roosevelt of the need for such an agency. Having observed the successes of the fascist fifth column in Europe, he urged the development of an international secret service for the United States to meet the Nazi challenge. His forceful advocacy of American involvement in the European conflict as well as his prediction that England would not collapse under the pounding of the Luftwaffe had impressed Roosevelt. Because of this and his personal audacity and imagination, he was Roosevelt's choice for this bold new venture.

One division of the COI was the propaganda wing, headed by playwright Robert E. Sherwood and staffed by such writers as Thornton Wilder and Stephen Vincent Benet. From the beginning there was friction between Donovan and the men he recruited—lawyers, bankers, PR men (all men of action)—and the sensitive writers.

Six months after the attack on Pearl Harbor (May 1942), the propaganda wing was split off from the COI to become the OWI, the Office of War Information, leaving the other activities to be directed by Donovan in a new agency, the Office of Strategic Services (OSS) with an ambiguous mandate "to plan and operate special services as may be directed by the United States Joint Chiefs of Staff." (Smith, 1972, p. 2)

The variety of activities in which the OSS engaged is suggested by the titles of its various branches: SI, Secret Intelligence; R & A, Research and Analysis; SO, Special Operations—destructive operations behind enemy lines working with resistance groups; MO, Morale Operations—black propaganda in contrast to the white propaganda of OWI; Counter-Intelligence; a Schools and Training Branch where recruits would learn the tricks of the OSS trades; and still others.

During the first year of its operation, there were three channels of entry into the OSS: recruitment of military personnel by the Personnel Procurement Branch, recruitment of civilians by the Civilian Personnel Branch, and recruitment of both military and civilian personnel through the initiative of individual OSS members—all of this without benefit of any professional or uniform screening process. Nobody knew who would make a good spy or an effective guerrilla fighter. Consequently, large numbers of misfits were recruited from the very beginning, and this

might have continued had it not been for several disastrous operations such as one in Italy for which, on the assumption that it takes dirty men to do dirty works, some OSS men were recruited directly from the ranks of Murder, Inc., and the Philadelphia Purple Gang. The need for professional assistance in selection was obvious, but was resisted by many in the organization.

In October 1943 an OSS official back from London suggested that a program of psychological/psychiatric assessment similar to that in the English War Office Selection Boards (WOSBs) be set up in the OSS. This idea was picked up and pushed by Robert C. Tyron, a psychologist on leave from the University of California, who was Deputy Chief, Planning Staff, OSS. He recommended that an assessment center be set up in the Schools and Training Branch. In collaboration with three other California Ph.D.'s, James A. Hamilton, John W. Gardner, and Joseph Gengerelli, he set about planning the first assessment center in the United States.

By November, a physical facility had been acquired: the Willard Estate in Fairfax, Virginia, some 18 miles from Washington. This estate, the spacious residence and grounds of the owners of the Hotel Willard in Washington, was to become Station S (for Schools and Training).

The first planning conference for the program of Station S was held in early December with Henry A. Murray (Harvard) and Donald Adams (Duke) joining the California psychologists. Shortly thereafter the director of the OSS authorized the establishment of an assessment unit, and 15 days later a skeleton staff held the first assessment program.

Although the push for an assessment program in the OSS came from the California psychologists, the ultimate form and nature of the program was shaped primarily by Henry A. Murray.

In addition to the speed with which it was initiated, the assessment program suffered from other handicaps. Although backed by General Donovan and some of the bureau chiefs, it was opposed by others, especially the military. We lacked knowledge about the assignments, most of them novel, to which our assesseees would be sent. Without job analyses, we did not know specifically for what we were assessing. We needed experts to write job descriptions, but there were none in the field. At best, job assignments were described by single terms: language expert, cartographer, news analyst. Had the OSS known what specific skills would be required, there would have been so many of them as to preclude testing them all.

Later we would learn more about what was required for successful execution of OSS assignments from branch chiefs who had by then received more specific job descriptions, from reports of returnees, from assessors who had received training in the OSS schools, and from assessors who had traveled abroad for firsthand observations.

In the beginning it was the lack of specific knowledge that led us to conclude that assessments could not be made of the specific skills of a given candidate for a specific job but rather in each case an assessment of the "man as a whole," the general structure of his being, and his strengths and weaknesses for rather generally described environments and situations. As it turned out, there were some advantages to our having taken this stance toward assessment, for we soon discovered that very often assesseees were never assigned to the job for which they were recruited. Typically, two to eight months elapsed between assessment and job assignment overseas. The candidate had to be trained first and by the time his training was completed, the war had moved on and the job for which he had been recruited no longer needed to be done.

Only those destined for overseas for assignment were assessed; those who remained in the United States were exempt. At first our assessment reports were only for the information of bureau chiefs. They were free to accept or reject recommendations as they saw fit. After two months, this policy changed. By order of General Donovan only those who received a positive recommendation from Station S could be sent overseas. This was personally flattering, but very frustrating to our scientific egos since it meant that anything like pure validity studies of our assessment operation could not be made. This order also meant that Station S with its 3½-day program would not be able to assess all those destined for overseas duty. Thus it was that in late winter 1944, a one-day assessment center, Station W, was set up in Washington to assess a large number of candidates, many of whom were to be assigned to headquarters and rear bases overseas rather than to operations in the field. Two months later a center to assess candidates recruited on the West Coast was established at Laguna Beach, California. This was Station WS. Later, assessment stations to screen native agents were set up in Ceylon, Kunming (Yunan Province), Calcutta, and Hsian. During the period of their operation, Stations S and W assessed 5,391 recruits.

The program for Station S, set up hurriedly and with little knowledge of what OSS assignments would entail, was bound to undergo many changes. There were seven periods in the history of Station S, but no radical changes in the program were made during the last six periods (June 1944 to V-J Day in September 1945), during which time I served as Director of Station S, and this the program I shall describe.

I have spoken of some handicaps under which the assessment program had to operate. Let me mention two others. From its inception,

the OSS had to guard against infiltrations by foreign agents and, of course, if its operations were to succeed, they had to be kept secret. But these concerns hardly justified the extreme secrecy which was maintained and which merely added to the atmosphere of cloak-and-dagger mystery that enshrouded the organization. This atmosphere was nowhere more obvious than in the recruitment of OSS personnel. Barred from mentioning the OSS by name but free to talk about mysterious, exciting overseas assignment with a government agency, the pitch made by OSS recruiters was especially attractive to the bored, to the pathologically adventuresome, to those neurotically attracted to danger, and to psychopaths in general. Thus we had more than our share of misfits to weed out, and of course it is psychopaths who have a special talent to make a good impression over brief periods of time. Under the given circumstances, the branch representatives who briefed candidates for their visit to Station S were wary about telling them much about the work for which they had been recruited. Many, when seen in assessment, had no idea of what their assignment would be, and some even thought they were in the State Department!

The other handicap we faced, or so it seemed at first, was that we would not know whom we were assessing. That is, they would not be allowed to reveal to us or to the other members of their assessment class their true names. In December 1943 when Station S was established, many of the recruits into OSS were refugees from Europe, often with families and relatives in occupied territory. If their true identities were to become known to others in the organization who had infiltrated it, they might become subject to blackmail through threats to relatives abroad.

Not knowing the identities of our assessee was not, however, entirely disadvantageous.

Quite innocently the operation could be a check on the misuse of nepotism, turning down even a Roosevelt or a Donovan if his performance at Station S left something to be desired.

In any case, we had to live with the anonymity of our assesseees so what we did was to turn it into an assessment exercise!

Assessment at Station S ran in two-week cycles, during which three groups of 18 candidates would be assessed. The first group would arrive on a Sunday afternoon and leave the following Thursday morning. Thursday afternoon a second group would arrive, were assessed over the weekend, and would leave Monday morning. Monday afternoon a third group would arrive and leave Friday morning. That weekend the staff would draw its breath and prepare for the next two-week cycle which would begin again on Sunday afternoon.

Upon arrival in Washington, each recruit for OSS reported to the Branch representative who would supervise him throughout training and who had authority then and there to reject the candidate if he were obviously unsuitable. Otherwise and usually, the recruit would be told something, but not much, about the type of work for which he had volunteered, and he would fill out a Student Information Sheet.

He was then informed that he would spend the next 3½ days at an assessment school in the country, where he would be asked many questions and given many tests by a group of psychologists and psychiatrists. He was told there was no reason to feel apprehensive, that most recruits who had gone through the school had enjoyed it. He was urged to do his best. He was also told he would go incognito for his own protection. Finally, he chose a name by which he would be known during assessment.

During that day in Washington he was given a medical examination, and that afternoon he would report to Schools and Training

Headquarters (an old brick school house that stood where Watergate now stands). He would report to headquarters at 3:30 p.m. if he was an enlisted man, at 4:00 p.m. if a civilian, and at 4:30 p.m. if an officer. Reporting was staggered in this fashion so that each might know the status of some of the men in his own group, the fewer the better, but not the others, e.g., no officer would know who were enlisted men and who were civilians, et cetera.

Upon arrival at Headquarters, the candidate would be taken to a room, stripped of his clothing and anything else that might identify him, issued army fatigues and boots, and ushered into a waiting area. At 5:00 p.m. the group would be put in a closed army van and driven over the 18 winding miles to Station S.

After a brief welcome and orientation to the program, the candidates were told that during their stay at S each would have to build up and maintain as completely as possible a cover story for himself, claiming to have been born where he wasn't, to have been educated in institution other than those he had attended, to have been engaged in work or profession not his own, and to live now in a place that was not his true residence. The cover story was to be maintained with staff and students alike at all times except under X conditions, at which time candidates could reveal anything about themselves except their names and true identities. X conditions prevailed when filling out a Personal History Form, in a Life History Interview, when a staff member established X conditions, and when in case of need a student asked for X conditions and a staff member granted them.

Thus Station S was one grand simulation, each candidate having the task of consistently simulating a person he was not. Since many of the candidates would eventually be living abroad under cover, the relevancy of the cover story task seemed clear enough.

The variables rated on the Station S report for each candidate offer a rationale for the assessment exercises which were developed for the program. The rated dimensions were: Motivation, Practical Intelligence, Emotional Stability, Social Relations, Leadership, Physical Ability, Observation and Reporting, Propaganda Skills, and Maintaining Cover. Each variable was rated on a six-point scale: Very Inferior, Inferior, Low Average, High Average, Superior, Very Superior. Note: To force a spread of ratings, using more of the whole scale, no rating of Average was given. If the dimension was not measured, that fact was indicated in the final report.

Components of each of the dimensions were printed on the face sheet in the following manner:

- > **Motivation:** energy, zest, effort, initiative, war morale, interest in assignment
- > **Practical Intelligence:** speed and accuracy of judgment, resourcefulness in solving problems
- > **Emotional Stability:** emotional control and maturity, absence of neurotic symptoms
- > **Social Relations:** social awareness, good will, teamwork, tact, absence of annoying traits
- > **Leadership:** social initiative, organizing ability, ability to evoke cooperation
- > **Physical Ability:** agility, daring, ruggedness, stamina
- > **Observation and Reporting:** ability to search, question, observe and recall, infer, report
- > **Propaganda Skills:** ability to affect others through acts, words, or displays
- > **Maintaining Cover:** caution, ability to remain inconspicuous, bluff, mislead, keep a secret

If one of the components of a variable was especially high (let us say tact), it would be underlined; if especially low, it would be crossed out.

The core of the assessment program was a rather detailed life-history psychiatric *Interview* conducted by a senior staff member. In preparation for the interview, the interviewer read a *Personal History Form* which the candidate had filled out, as well as a *Projective Questionnaire*, a *Sentence Completion Test*, a *Health Questionnaire*, and a *Work Conditions Survey*. The *Work Conditions Survey* described 43 conditions that the candidate rated on a 6-point scale ranging from (1) "Highly desirable" to (6) "Would make job impossible." In light of all this material, the interviewer had some impression of the candidate and knew the areas in which he would want to obtain more information in the interview.

Two conventional measures of intellectual ability were administered, the *Otis Self-administering Test of Ability* and a *Vocabulary Test*, plus two tests of special aptitudes, the *Bennett Mechanical Comprehension Test* and the *Signal Corps Code Aptitude Test*, but we were less interested in our candidates' scores on intelligence tests than in the effectiveness with which they used whatever intelligence they had. And this, as well as most of the dimensions of behavior which we rated, was more clearly to be seen in the most novel aspect of the OSS program—the situational tests which we developed, tests which today would be more often referred to as simulations or simulation exercises. Our objective was to observe as much of the candidate's behavior in situations which simulated as realistically as feasible the kinds of situations they would be likely to encounter in OSS assignments.

For those destined to work in Secret Intelligence, the abilities to search, question, observe and recall, infer, and report would be of crucial importance. A variety of tests (some situational) to tap the dimension called Observation and Reporting were devised.

Examples of such tests in which, of course, other variables also could be observed and rated were:

- > **Belongings Test:** In this test the candidate would be taken to a bedroom in which 26 items such as articles of clothing, written materials, a timetable, newspaper clippings, a ticket receipt, etc., were placed openly on the bed, chairs, and tables. His task was to examine them, to size up the man who left them, to learn all he could about the person (what he was like, etc.). Objects could be picked up and examined, but had to be put back in place. After four minutes the candidate was taken to another room to answer a 36-item questionnaire, a test of the candidate's ability to observe and draw correct conclusions.
- > **Map Memory Test:** The candidate is to assume that he is an agent operating in the field and that he has just made a secret rendezvous with a courier who has a map of the territory that the agent will be covering. After a few minutes the courier must leave with the map and since it would be dangerous for the agent to have it, he must memorize it. After eight minutes to examine the map, it is taken away and the candidate answers a set of multiple-choice statements about the terrain shown on the map.

Interrogation Test: In this situational test a junior staff member plays the role (a standardized one) of an escaped prisoner of war who is interrogated by the candidate. Background material for the interrogation is supplied in the form of a map and some general information about the imprisonment and escape of the prisoner who had served as a tail gunner of a B-29 before its crash landing. Another staff member sits by and rates the candidate's skill in Developing Rapport, Asking Productive Questions, Following Up Leads and Hints by the "Prisoner," etc.

For those destined to work in Morale Operations (MO), the ability to affect others through acts, words, or displays would be crucial, and special tests were devised in which these skills might be displayed, thus providing the assessors with some estimate of the Propaganda Skills of candidates. Two examples of this type of test of propaganda skills would be the *OWI Test* and the *Manchuria Test*. In the former, a test of one's sensitivity to cultural differences, the candidate was to assume that he was working for OWI in Korea and knew nothing about the country. He had 20 minutes to indicate the kind of information he would want to have in order to work up a propaganda program designed to win Koreans to our side. In the *Manchuria Test* the candidate was given some background facts and was, over a day and a half, to prepare two pieces of propaganda to lower the morale of Japanese railway workers and guards on the South Manchurian Railway. In this task, candidates were urged to use whatever graphic skills they had.

One of the better known of our situational tests was the *Brook Test* (a Leaderless Group Situation), in which a group of six candidates would be taken to a shallow, quiet stream whose banks were eight feet apart. On one bank was a heavy rock; on the other, a log. There were trees on both sides and scattered on the bank where the group stood were a number of boards (none long enough to reach from bank to bank), three lengths of rope, a pulley, and a barrel with both ends knocked out.

The candidates were told, "In this problem you have to use your imagination. Before you, you see a raging torrent so deep and so fast that it is quite impossible to rest anything upon the bottom of the stream. The banks are sheer, so it will be impossible to work except from the top of them."

“You are on a mission in the field and having come to this brook, you are faced with the task of transporting this delicate range finder, skillfully camouflaged as a log, to the far bank and of bringing that box of percussion caps, camouflaged as a rock, to this side. In carrying out this assignment, you may make use of any materials around here. When the job is done, all of you, as well as any material you used, are to be back on this side.” (OSS Assessment Staff, 1948, pp. 95-96)

The group was to work in an area, marked by two white stakes, about 15 feet along the bank.

In this exercise one could easily note the difference between asserted leadership and effective leadership, and ample opportunity was given to observe differences among the candidates on such variables as Energy and Initiative, Effective Intelligence, Social Relations, Leadership, and Physical Ability.

A second Leaderless Group Situation was *The Wall*. Here the candidates were led to a wall 10' high, 15' long. Behind the wall which the candidates could see was another wall of the same height and length, parallel to it, the two walls being 8' apart. On the candidates' side of the wall there were a heavy log, an old board a few inches longer than the log, and a couple of two-by-fours, two feet and three feet in length.

The candidates were informed (fancy being mixed with fact as in all situational tests) that although they could not see it, the barrier before them actually consisted of two walls (fact) separated by a 200-foot canyon (fancy). They were escaping from some Japanese soldiers. In order to save themselves they would have to get to the other side of the far wall and to fulfill their mission they would have to take their king-size bazooka (the log) with them. They might get across the wall in any way they wished, but they must not walk around the ends of the wall and, of course, whoever or whatever fell into the “canyon” would be counted as lost.

The same variables observed and rated at *The Brook* were observed and rated at *The Wall*.

The Brook and *The Wall* gave some advantage to men with marked initiative and assertiveness. To discover real ability when leadership was assigned, five other field tests were employed. The task of the leader in the *Mined Road* was to get his group across a mined road in 10 minutes; in the *Sentry*, to get his group, with one man wounded and unable to walk, across a road guarded by a sentry (played by a junior staff member in uniform) whose regular beat meant that for about 45 seconds of each of his tours he could not see the section of the road where the group would try to cross. In the *Agent's Room*, the leader has just received a radio message that a friendly agent had, in a hurried getaway, left some important papers behind; the task would be to recover the papers. The instructions were given at the entrance to a large, unused carriage house on the second floor of which was a crude living room (the agent's room). Members of the group were told that they had been passing as natives, had forged credentials, were dressed in native costume, and could speak the native tongue. Midway through this exercise a German guard armed with a rifle approached the carriage house.

The other two assigned leadership exercises, the *Blown Bridge* and *Killing the Mayor*, like the three just described, provided good opportunity to observe each candidate's leadership skills as well as other variables: Practical Intelligence, Emotional Stability, Social Relations, and Physical Ability.

An indoor, verbal counterpart of the outdoor, physical *Brook* and *Wall* exercises was a Leaderless Group Discussion. The topic was, “What are the major postwar problems facing the United States and (if you have time) along what lines do you think they should be solved?” The dimensions observed and rated

in this situation were Leadership, Effective Intelligence, Energy and Initiative, and Social Relations.

To measure resistance to stress and frustration tolerance, important facets of Emotional Stability, two situational tests were devised. *Construction*, which came to be known as *Behind the Barn*, for that was where this exercise was held, required the candidate to direct two helpers in the task of building a five-foot cube structure with seven-foot diagonals on the four sides, using a huge “tinker toy” set of materials. The candidate had 10 minutes in which to accomplish the task. All the physical work was to be done by the helpers, junior staff members who played the role of Kippy (passive, sluggish, and something of a stumblebum) and Buster (aggressive, critical, constantly making impractical suggestions). Both were insulting, faultfinding characters. In the history of Station S, this job was never completed in the allotted time. Some candidates gained insight into the problem, but more often they became so involved and so frustrated that they had difficulty in handling their frustration and controlling their anger. A few physically attacked their helpers, and some asked to be relieved from the program after this exercise.

After the *Construction* test, the candidate was taken aside by a senior staff member who in a *Post-construction Interview* provided the candidate an opportunity for catharsis. Some admitted their anger and confessed to having almost hit their “helpers” while others, with quivering lips and trembling hands, denied they had been in any way emotionally upset in the exercise. This interview was also used to test the candidate’s ability to maintain cover. Since a person is particularly inclined to break cover in moments of relaxation following stress, the interviewer would inquire about

earlier similar experiences which the candidate had had and how he had acted then. If the candidate was not on guard, he would be led into talking about past experiences of a similar kind in which he had performed better and this, of course, was a breaking of cover since X conditions had not been established.

In the *Stress Interview* proper, candidates, one at a time, were instructed that they have just been found by a night watchman in a Government building in Washington going through papers in an office. The candidate is not an employee of the agency, nor does he have any identification papers. The night watchman is bringing him to the security office for questioning. The candidate was given 12 minutes in which to invent an innocent cover story that would explain his being in the building, after which he was to knock on designated basement door. Upon being ordered to enter, he was seated in a dark room, the only light being a blinding spotlight directed at his face. For 10 minutes he would be questioned about his presence in the building. At first quietly and in an understanding, sympathetic manner, he was encouraged to give more and more detail, then suddenly the questioner would turn sadistic, confronting him with inconsistencies and implausibilities in the story with an ever-increasing tempo and crescendo of voice. The staff member who questioned the candidate was a New York trial lawyer skilled in techniques of cross-examination. Suddenly the examiner would break in with, “We now have abundant evidence you have not been telling the truth. That is all.” Then after whispered conversation among the staff who had been sitting behind the examiner observing the candidate, but not seen by him except as dim forms, the examiner would announce, “It is our decision that you have failed this test.” After five seconds, during

which his reactions were observed, the candidate was told to go to a room on the third floor. There he found a staff member, feet on the desk, reading a paper, who would inquire, "How are things going?" In this *Post-stress Interview*, many a candidate would break cover when sympathetically asked if he had ever been in a comparable situation. *The Stress and Post-Stress Interviews* provided ample opportunity to observe and to rate Emotional Stability, Maintaining Cover, and Motivation for assignment in OSS.

Late in the afternoon of the second day all candidates under 45 years of age who had passed the physical examination before coming to Station S, ran an *Obstacle Course* especially designed to test physical agility and daring.

That evening, candidates were assigned in pairs to act in *Improvisations*. These were role-playing situations dreamed up by the staff in sessions referred to as "brain storms" and tailor-made for each candidate to help resolve doubts which the staff might still have about him. For example, there might be some question left as to how Candidate A takes criticism or rejection and how much tact Candidate B can muster. The situations they might be called upon to improvise, with emphasis put on what they would actually do if confronted with such a situation rather than how they would act it out, could be the following:

"A moved to a small city about three months ago and opened a business there. He has been doing quite well and one month ago sent in an application for membership in a club in the town. He has heard nothing in response to his application and goes to the home of B, a prominent member of the club, with whom he is pleasantly acquainted. (A is then sent out of the room and B is told that A has received several blackballs.) A is then called back into the room (OSS Assessment Staff, 1948, p. 170) and the improvisation begins."

After each improvisation the participants' reactions were questioned, commented upon, and criticized by the other participants, and also by the other candidates and staff. Each candidate was observed in two improvisations. At a halfway break, liquor was brought in and continued to be available through the evening and as long into the morning as anyone wanted.

After *Improvisations*, candidates were divided into two teams to debate the question, "What Shall We Do with Germany After the War?"

The Debate provided opportunity to observe and rate the dimensions of Energy and Initiative, Effective Intelligence, Social Relations, Leadership, and (especially since liquor was still flowing) the Ability to Maintain Cover.

The unmentioned and never named *Liquor Test* led to much behavior which otherwise would not have been observed. During the last months of OSS we were grateful for an OSS operator (an Army Captain who had parachuted into France to work with a resistance group) who, staying up as late as any of the candidates wished, relieved the professional staff of this duty.

On the morning of the last full day of assessment, the candidate's opinions of and attitudes toward each other were collected. The *Sociometric Questionnaire* asked candidates, "With whom would you enjoy continuing your acquaintance?" "Which men expressed the most realistic and convincing opinions in the debate last night?" "If you were a member of a group on a dangerous mission, whom would you prefer to have as your leader?" and other questions of the same sort. *The Judgment of Others* test required the candidates to write five personality sketches for the five assesseses whom they felt they had come to know best. Data from these two exercises were available to the staff to aid in forming their final impressions of the candidates.

During the rest of the day, while the staff was preparing its S Report on each of the candidates, the assesseees were involved in solving the *Murder Mystery*. For this exercise they were given copies of the mythical *Fairfield Chronicle* which reported the finding of the dead body of a woman on a road near S. Working in competing groups of six, their task was to wring from three junior staff members (standard role players) who were designated as possible informants the information necessary to make proper inferences and thus solve the murder.

Athletic Events—broad jump, high jump, and shot-put—more in the spirit of a game than as a serious test, and a *Baseball Game* between the candidates and junior staff members not involved in the staff conference engaged the interest of the assesseees during the last afternoon.

Reports on the candidates' performance in these last day simulations and games were available to the staff before their deliberations ended.

For each assessment class, usually consisting of 18 assesseees, the staff was divided into teams of two senior staff members (professionals with Ph.D. or M.D. degrees) and one junior staff member (enlisted men who had had some training in psychology). Each team was assigned to a group of five to seven candidates. The senior members conducted the life history interviews; the junior member administered special individual tests and interpreted the projective test protocols. Otherwise, the work of both senior and junior members was the same, namely, to develop as a group as complete a conception as possible for each candidate in the subgroup assigned to them.

During the various situational tests the behaviors of each participant were carefully noted by both senior and junior members of the

responsible team, each staff member rating each assessee on the variables relevant to the particular exercise. Usually immediately following each situational test, the staff team met to discuss their impressions of the candidates and the ratings they had assigned to them. The purpose of the meeting was to come to agreement upon the ratings to be assigned to each assessee on each of the rated variables.

Staff meetings for the preparation of the Station S Report on the candidates ran from after lunch on the last full day of the assessment until all reports had been finished, which was often close to midnight.

Before the staff meeting each interviewer had written a first draft of his section of the reports on the men he had interviewed and for whom he was the primary assessor. This was a description of the candidate as a person and of the major events of his past history that seemed to shed light on his present personality, with special emphasis being given to picturing how the candidate might be expected to behave in a variety of situations in the future. Also before the staff meeting, the situationist (the member of the staff who was responsible for reporting on the behaviors of the candidate at Station S as revealed in the various situational tests) had written his section of the final report, which not only described the behavior of the candidate at S but also offered specific support for the statements and predictions made in the first part of the report written by the interviewer.

A huge board which covered one wall of the staff room offered a graphic display of how each candidate had been rated by the staff team assigned to him on each of the variables in each of the relevant tests and situations. Ratings were represented by thumbtacks; red if above average, blue if below. While the

interviewer and situationist read their reports, staff members looked at the board and the thumbtacks to see if the reporters' statements were supported by the ratings. If not, there was discussion until general agreement was reached about the changes that would have to be made in the overall rating of the variables on the Station S Report or in the statements in the written report. Sometime the reports could be revised in the staff meeting. If not, they would be rewritten by the report writer after the meeting.

The final report consisted of a face sheet on which the candidate was rated on the nine variables already described, other sheets which carried a character sketch of the candidate based upon the interviewer's insights and the staff's observations of the candidate at S, and recommendations concerning overseas assignment made on a five-point scale: Not Recommended, Doubtful, Recommended with Qualifications, Recommended, Highly Recommended. A similar scale (Not Recommended to Highly Recommended) was used to indicate the candidate's fitness for work at (1) a rear base, (2) an advanced base, and (3) at or behind enemy lines; his fitness for (1) high, (2) middle, or (3) low level of authority and responsibility; and finally his fitness for different types of assignment, e.g., administrator, intelligence officer, operational agent, etc.

These Station S reports, sometimes worked over long past midnight of the last day of assessment, were already at OSS Headquarters in Washington when the candidates arrived back there the next morning.

How effective was the OSS assessment program? We cannot say with certainty; but accepting our validity figures at face value, we were forced to conclude that we were not very successful in predicting performance overseas. But were errors mainly in the

assessment process or in the appraisal process, or in both? Again we cannot say with certainty, but we do know that the appraisal process as carried out left very much to be desired.

Our appraisal data were of four types:

1. **Overseas Staff Appraisal:** These were appraisals made by OSS assessment staff members on the basis of interviews with the immediate chief or commanding officer and, if possible, associates of our "graduates" overseas.
2. **Theater Commander's Appraisal:** From Spring 1944, each returnee was reported on and his personality traits were rated by his immediate superior.
3. **Reassignment Area Appraisal:** From Fall of 1944, a center was set up (Area F) for the reassessment or re-evaluation of personnel who had completed a tour of duty in Europe or the Middle East for reassignment in the Far East.
4. **Returnee Appraisal:** Beginning in the Summer of 1945, each returnee was asked to report on others known to him, rating their performance Low, Medium, or High.

When assessment job ratings were correlated with appraisal job ratings for Stations S and W, validity coefficients were all positive but disappointingly low, ranging from .08 to .53, depending upon the source of the appraisal data, as shown in Table 1. Much to our surprise the one-day assessments at W appeared to have been more effective than the 3½-day assessments at S.

We of the OSS staff are indebted to Jerry S. Wiggins (1973) who, making certain assumptions and using improved principles for estimating outcomes of predictions, worked over our data and came up with a more favorable picture than the one we had drawn.

**Correlations between S and W Assessment Job Ratings and Appraisal Ratings
(After OSS Assessment Staff, 1948, p. 423)**

Type of Appraisal	S Job Rating (Classes S-45 on)		W Job Rating (All Classes)	
	r	N	r	N
Overseas Staff Appraisal	.37 ^a	88	.53 ^a	83
Returnee Appraisal	.19 ^a	93	.21 ^a	173
Theater Commander's Appraisal	.23	64	.15	158
Reassignment Area Appraisal	.08	53	.30 ^a	178

^a Cases in which correcting r for restricted sample made a significant difference; r given in each case is the corrected one.

Table 1

His estimate is that at S, if we had used only random selection, our percentage of correct decisions would have been 63%, but actually 77% were correct. Corresponding estimates for Station W are that by random selection 66% would have been correct, but actually 84% were correct. This means that at Station S, assessment effected a 14% increase in correct decisions over random selection and at Station W an 18% increment. Considering the crucial nature of the assignments, increments of 14% and 18% of correct decisions are not unimportant.

It is interesting to speculate as to why the briefer (1 day) assessments of Station W were more accurate than the longer (3½-day) assessments of Station S. It may have been that the procedures used at W were more efficient than those employed at S, although this seems unlikely. Perhaps the staff at W was more competent than the one at S: more psychiatrists served on the staff at W and less use was made of junior assessors than at S. Differences in the populations assessed at the two stations could also have been a determining

factor. Those assessed at W were more often high echelon executives in the organization, women secretaries, and office workers, many of whom had already spent some time in the Washington headquarters; while those sent to S for assessment were either the more difficult cases who were already presenting perplexing problems or were men destined for more hazardous duty under more stress and danger than would be experienced by the W assesseees. Indeed, of those assessed at W, 74% received rear base assignments with only 15% serving behind enemy lines. In contrast, only 29% of Station S graduates received rear base assignments while 43% operated behind enemy lines. Among possible explanations of the differential success rates of the two assessment centers, the one that seems least plausible is the notion that the staff at Station S suffered from a superabundance of information about their assesseees, while the staff at Station W with less information had just what they needed to make the kinds of decisions called for in the OSS assessment program. But the fact remains that we cannot say with certainty why the assessments at W surpassed

those at S. Indeed, still today the optimal length of assessment center programs remains an unanswered question, one which should have been subjected to empirical investigation long ago.

At the end of their report on the OSS program, *Assessment of Men*, the OSS staff made a number of recommendations which they hoped would remedy some of the defects of assessment programs as practiced in the OSS.

These recommendations were published 26 years ago, 10 years before the first operational assessment center was established in American industry by Michigan Bell. The recommendations were formulated as definite rules although we recognized that they were no more than a set of hypotheses to be tested in the planning and operating of subsequent assessment centers which it was our hope would be established.

These recommendations are reproduced below with the thought that the readers of this monograph may find it both interesting and informative as they review them to ask themselves such questions as these: How many of these recommendations have been carried out in the setting up of assessment centers in business, industry, government, and education? If adopted, have they proved helpful? And if they have been ignored, were they rejected for good reason? And finally, what further recommendations should now be made for the improvement of assessment centers.

The recommendations follow:

1. Select a staff of suitable size and competence, diversified in respect to age, sex, social status, temperament, major sentiments, and specific skills but uniform in respect to a high degree of intellectual and emotional flexibility (p. 473).

2. Before designing the program of assessment procedures, conduct a preliminary study of the jobs and job holders of the organization (p. 475).
 - 2.1. Make an adequate functional analysis of each of the roles for which candidates are to be assessed, as well as an analysis of the environments in which each role must be fulfilled (p. 476).
 - 2.2. Obtain from members of the organization a list of personality attributes which, in their opinion, contribute to success or failure in the performance of each role (p. 476).
 - 2.3. After a careful survey, analysis, and classification of the information obtained by these observations and interviews (recommendations 2.1 and 2.2), make a tentative list of the personality determinants of success or failure in the performance of each role. These determinants will constitute the variables which, if possible, will be measured by the assessment procedures (p. 477).
 - 2.4. Define, in words that are intelligible to members of the organization, a tentative rating scale for each personality variable on the selected list as well as for the overall variable, Job Fitness (p. 479).
 - 2.5. Devise a satisfactory system for appraising the performance of members of the organization both at this time and later (p. 481).
 - 2.6. Obtain appraisals of a properly distributed sample of the present members of the organization (p. 484).
 - 2.7. Examine the defects of the appraisal system as revealed in practice (recommendation 2.6), and correct these by revising, where necessary, the lists of variables, definitions, rating scales, or other elements.

-
- 2.8. Obtain the figures necessary for a brief numerical statement of the personnel history of the organization over the last four or five years (p. 485).
 3. Design a program of assessment procedures which will reveal the strength of the selected variables; for assessing these variables, set up scales which conform to the rating scales that were defined for the purpose of appraisal (p. 485).
 4. Build a conceptual scheme in terms of which formulations of different personalities can be made (p. 488).
 5. Set up an efficient punch-card system which will permit periodic statistical analyses of assessment findings (p. 490).
 6. Assess candidates for a long trial period without reporting ratings for decisions to the organizations (p. 491).

How far and in what directions the state of the art of assessment has moved beyond that obtained in the assessment program of the OSS is a fascinating chronicle, but that is another story for another time.

REFERENCES

OSS Assessment Staff. (1948). *Assessment of men*. New York: Rinehart and Company.

Smith, R. H. (1972). *OSS: The secret history of America's first central intelligence agency*. Berkeley, CA: University of California Press.

Wiggins, J. S. (1973). *Personality and prediction: Principles of personality assessment*. Reading, MA: Addison-Wesley Publishing Company.

ABOUT THE AUTHOR

Donald W. MacKinnon

The late Donald W. MacKinnon received his undergraduate training at Bowdoin College and did graduate work at Harvard University, receiving a Ph.D. in psychology from Harvard in 1933. He taught at the University of Maine, Harvard University, Radcliffe College, Bryn Mawr College, and, from 1947 until his retirement in 1970, at the University of California, Berkeley.

From 1930-1931 he studied in Europe as a Sheldon Traveling Fellow from Harvard University. In the summer of 1953 he taught at the Salzburg Seminar in American Studies in Salzburg, Austria. He also held summer session teaching positions at the University of Minnesota (1940) and at the University of Hawaii (1955 and 1969).

During World War II Dr. MacKinnon was a member of the assessment board and, from June 1944 to August 1945, he was director of Station S in the Office of Strategic Services (OSS), the precursor to the Central Intelligence Agency.

From 1947 until his retirement in 1970, Dr. MacKinnon served as director of the Institute of Personality Assessment and Research on the UC-Berkeley campus.

Dr. MacKinnon was the author of several scientific articles and co-author of two books: *Assessment of Men*, a report of the OSS's assessment program during World War II, and *Experimental Studies in Psychodynamics*.

Dr. MacKinnon was the American Psychological Association's 1962 Walter Van Dyke Bingham Memorial Lecturer. His lecture, "The Nature and Nurture of Creative Talent," given at Yale University in April 1962, was published by *The American Psychologist*. In 1967 he received the American Psychological Association's Richardson Creativity Award in recognition of his research in creativity.

During 1973-1974, Dr. MacKinnon was Visiting Fellow at the Center for Creative Leadership, Greensboro, North Carolina. During that time he made an extensive study for the Center of the current status of assessment centers in the United States.

Dr. MacKinnon died in 1987.

THE AMERICAS

WORLD HEADQUARTERS
PITTSBURGH
412.257.0600

MEXICO CITY
52.55.1253.9000

TORONTO
416.601.5500

Other major offices in
Atlanta, Chicago, Dallas,
Detroit, Los Angeles,
Monterrey, Montreal,
New York, St. Louis,
San Francisco, Santiago,
and São Paulo

EUROPE/AFRICA

DÜSSELDORF
49.2159.91680

LONDON
44.1.753.616.000

PARIS
33.1.41.9686.86

Other major offices in
Johannesburg, Utrecht,
and Warsaw

ASIA-PACIFIC

HONG KONG
852.2526.1188

SINGAPORE
65.6226.5335

SYDNEY
61.2.9466.0300

Other major offices in
Auckland, Bangkok,
Jakarta, Kuala Lumpur,
Manila, Melbourne, Seoul,
Shanghai, Taipei, and Tokyo

E-MAIL: INFO@DDIWORLD.COM
WWW.DDIWORLD.COM



MXPH

MKTCPMN07-1294-0105

ABOUT DDI. Since 1970 Development

Dimensions International has worked with some of the world's most successful organizations to achieve superior business results by building engaged, high-performing workforces.

We excel in two major areas. Designing and implementing selection systems that enable you to hire better people faster. And identifying and developing exceptional leadership talent crucial to creating a workforce that drives sustained success.

What sets DDI apart is realization. We focus on the needs of our clients and have a passion for their success.

The outcome? You bring the best people on board, who get up to speed faster, contribute more, and stay longer—giving you the ultimate competitive advantage.



POWERING
SELECTION
SUCCESS



DEVELOPING
EXTRAORDINARY
LEADERS



UNLEASHING
EXECUTIVE
TALENT

