

*Hugh Murray*

---

## The Transformation of Selection Procedures

### The War Office Selection Boards\*

#### *The Presenting Problem and the Initial Response*

Towards the end of 1941, the impending rapid expansion of the British Army required a large number of officers. The ensuing crisis in officer selection was of sufficient magnitude for a major innovation in assessment procedures to emerge—the War Office Selection Boards. These boards enabled the army to officer itself when traditional methods were failing and when there was doubt as to whether a sufficient reserve of officer material existed among the other ranks. The process of collaboration between experts and administrators, which the boards exemplified, became a model for many other joint undertakings. The methodological revolution consisted in replacing a military judge using a short interview by an inter-disciplinary group of selectors who assessed groups of candidates over two-and-a-half days. The extent of the participation achieved among all those concerned made the Boards profoundly acceptable to the war-time army.

Failure rates at Officer Cadet Training Units (OCTUs) had risen to over 20 percent in many courses and to over 30 percent in some. Not only did these failures represent a great deal of effort wasted on unproductive training (courses were of three months), they created undue stress in the training units. Next, there were insufficient numbers of good applicants. This lack was complex in its origins; letters of complaint received by the War Office indicated that there was a reluctance to apply for a commission. Furthermore, the return to their units of a noticeable number of failures reinforced this reluctance.

At this time candidates for commissions went before a Command Interview Board (CIB) consisting of a permanent president and two commanding officers (COs) as ad hoc members. These boards conducted a short interview with each

\*A new paper based on original documents and unpublished papers, drafted by B.S. Morris, J.D. Sutherland and E.L. Trist, held in the archives of the Tavistock Institute.

candidate, usually some 20 minutes, and decisions were made on the impressions gained, together with the information contained in a brief report from the man's CO. Many candidates felt that they could not do themselves justice in such short interviews. The presidents felt equally dissatisfied. Reports from units were also proving less helpful than had been hoped. COs had not yet had experience of men under sufficiently varied conditions and the course of the war had been such that relatively few candidates had had the critical test of battle. Potential officers were being drawn from an ever-widening range of social classes so that presidents no longer had those signposts to leadership qualities with which they were familiar in young men from the public (USA = private) schools. The uncertainty felt about such short interviews was increased by pressure to find all possible candidates rather than to take only those who were obviously good.

#### BACKGROUND TO THE CHANGE

Early in 1940 a psychiatrist was posted to each of the Army Commands and soon afterwards other psychiatrists were added to assist the Command psychiatrists. Many of the breakdowns they encountered were obviously precipitated by factors in the military environment as well as by limitations in the individual. The psychiatrists began to occupy a therapeutic role in relation to their employing institution, the army, as well as to individual patients within it by making suggestions for the prevention of psychiatric illness from social causes (Sutherland and Fitzpatrick, 1945). One of the most important causes of difficulty in adjustment was unsuitable employment in the army itself. A new Directorate for the Selection of Personnel was established which, working in close collaboration with the Directorate for Army Psychiatry, prepared a scheme which radically altered the recruiting arrangements of the army and entailed the building of a new social system—the General Service Corps—into which men were now taken for a short period before being sent to a specific arm of the service. During this induction period they were given several psychological tests and a short interview which enabled Personnel Selection Officers to make recommendations for each man's training in keeping with his abilities and, as far as possible, his preferences (Vernon and Parry, 1949). The psychometric underpinning of this scheme was in sharp contrast to the methods then used for the selection of officer cadets.

In creating practical schemes for handling various manpower questions, it became the rule that the schemes had to be jointly planned by the army officer and the "expert," each contributing from his own special experience and knowledge (Rees, 1945). The social-therapeutic role of the psychiatrists, both in diagnosing problems from the human side of the military environment and in

fostering the development of specially adapted military institutions to meet them, paved the way for the early experiments and for some of the most characteristic features of the selection boards.

#### PRELIMINARY EXPERIMENTS

Experiments by army psychiatrists with both officers and officer-cadets pointed to ways of providing the CIBs with more evidence than they were accustomed to have. An experiment by Bowlby was stimulated by comments on the unsuitability of many officers recently commissioned: the numbers of unsuitable officer-cadets were intolerably high, they lacked the ability to master the technical training or the degree of leadership required for an armored regiment, or both. Since psychological tests and interviews had proved useful in identifying other-rank recruits likely to prove failures, the new inquiry aimed to discover whether an intelligence test and an interview by a psychiatrist could accurately predict the technical ability and officer-like qualities of cadets at OCTU.

A critical experiment with serving officers (Wittkower and Rodger, 1941) arose out of a Command psychiatrist's work with problem officers and the interest of his Army Commander in the methods of officer selection used in the German army. An initial experiment was set up at a school for company commanders whose commandant and staff had, during an intensive five-week course, formed a thorough-going opinion of the students' all-around capabilities as officers, their technical proficiency and their human qualities. They could give ratings of their students with which the opinions of the psychiatrist could be compared.

The investigation included written and laboratory tests and an interview. In order to compare opinions, the commandant and the psychiatrists each made a brief evaluation of the student's personality, together with a judgment on his suitability as a combatant officer. Both sides read their reports and then rated the results of the comparison according to the degree of agreement. Of the 48 comparisons, 26 (55 percent) were in essential agreement, 12 (25 percent) in substantial agreement and 10 (20 percent) in essential disagreement. Nine-tenths of these disagreements were due to underlying personality deviations which had escaped the attention of the commandant, in some cases the psychological abnormality being very severe.

The program, with the addition of a psychologist to administer intelligence tests, was repeated with another course of officers at the company commanders' school. The results reduced the 20 percent disagreement by half. The improvement appeared to be due not only to the psychiatrists incorporating the results of the intelligence tests into their reports, but also to the mutual

education of the psychiatrists and the commandant. The psychiatrists learned more of the variety of talents which could successfully be used in officer roles, while the commandant became aware of the possible psychological significance of certain aspects of a man's performance during the course. It was recognized that some differences of opinion would be inevitable because of the limitations of the methods used by each judge—the one using interviews, supported by written and laboratory tests, and the other observing men in a variety of practical training activities.

The assessment of these officers, while presenting many difficulties, was nevertheless an easier task than the assessment of younger men who did not have occupational or military records as evidence of their potentialities. The investigations were therefore repeated with several groups of cadets at an OCTU. A similar degree of overall essential and substantial agreement, 80 to 90 percent, was found between the reports of the training staff and the psychiatrists. The relationship between training outcome and performance in the field was, of course, unknown.

So long as the categories of substantial and essential agreement were combined, the level of agreement was higher than might have been expected. But if the category of substantial agreement was added to that of essential disagreement, a more negative picture emerged. It was concluded that an opinion based simply on interview and intelligence tests would not be sufficient for making reliable judgments on the substantial proportion of candidates likely to be near the threshold of acceptance—and boards were under pressure to accept as many of these as they could, with safety, pass. From the nature of the discrepancies between the judgments of the psychiatrist and the OCTU staff, it appeared that practical tests would be a valuable addition to an interview, as a man could then be seen in action. If a way could be found of combining the resources and methods of military personnel and the opinions of psychiatric and psychological specialists, rather than of using one as a criterion for validating the other, a type of selection procedure might be instituted which would be reliable in assessing officer candidates and acceptable to military opinion.

#### INITIAL WORKING PRINCIPLES

From the preliminary experiments six general principles emerged for an improved selection procedure:

- The responsibility for selection must belong, and manifestly so, to the employing institution, i.e., the army. If selection were to be delegated to the "expert," insuperable difficulties would follow regarding the acceptance of new methods by both officers and men.

- The introduction of scientific procedures could best be effected by grafting them onto the existing Command Interview Boards. To do this entailed creating a new social institution, for the original board would be transformed in character. The president of the new board would retain responsibility for selection, but with evidence provided by other examiners. His experience of the army was essential. He should carry a rank—full Colonel—which would strengthen his position in relation to the COs of units from which candidates would be drawn.
- Data from interviews needed to be supplemented by observation of the individual in action. The president should have a junior regimental officer with experience of battle conditions, to be known as a Military Testing Officer (MTO), who would conduct a number of practical tests based on common tasks of an officer's role.
- The psychological contributions to the board's evidence should be of two kinds. First, evidence about each candidate submitted by a psychiatrist and a psychologist. As full an interview as possible should be preserved. In addition to tests which had been proved worthwhile, such as tests of intelligence, ways of estimating qualities of personality should be developed.
- Candidates should live in a hostel with the MTO for a period of three days—the time estimated to process an intake of, say, 30 candidates.
- Working out a practical testing procedure required further experimentation. An experimental board should, therefore, be established with a president, an MTO, two psychiatrists and a psychologist.

The proposals were well received. The creation of a new type of military unit for the selection of officers which would introduce scientific methods in the context of a residential procedure was acceptable to the presidents of the Command Interview Boards.

### *The Work of the Experimental Selection Board*

The experimental selection board\* assembled early in January 1942 to begin working out an operational procedure on the principles agreed. Whatever aspects of a candidate might need special attention in the light of job analyses—which were carried out by officers with recent battle experience—the

\*W.R. Bion, J.D. Sutherland (psychiatrists) and E.L. Trist (psychologist) became the nuclear technical group which, in conjunction with Colonel J.V. Delahaye (President) and Captain W.N. Gray (MTO), worked up and tested out a reproducible model. Three psychological assistants (sergeant testers—later commissioned) supported this group.

“whole man” had to be taken into account. Many kinds of men made good officers. Few personal qualities were specific to the job. Almost all an individual’s attributes could contribute to his effectiveness and could affect the attitudes of his men and of his fellow officers towards him. It would be his competence to fill the main roles of the officer’s job that would matter rather than his particular method of carrying them out. Preconceptions about officer qualities or types of potential officers had to be overcome. Judges needed an extensive knowledge of officer roles and then had to assess how candidates could use their resources to fill them.

Three main demands of the officer’s job needed assessment: quality of social relations with superiors, equals and subordinates; competence in practical situations; stamina over long periods and under stress. The president and the psychiatrist had their own distinctive method of assessment (the interview) already available but suitable testing methods for the MTO and the psychologist would have to be created.

#### QUASI REAL-LIFE SITUATIONS

The first military tests were decided by the background, training and battle experience which the MTO brought to his task. As a regimental officer he judged men on the basis of their performance in actual situations and roles. Therefore his intuitions and discriminations were likely to be most effective with tests which enabled him to relate what he observed directly to his field of experience. The most suitable tests were, therefore, quasi real-life situations in which the essentials of various officer roles and problems were imitated. The situations had to be such that they depended as little as possible on special military knowledge and amount of military training.

The tests were of two types: command situations and practical individual situations. Command situations consisted of asking each candidate to play the role of officer in simple military situations using the other candidates as his men. Such situations typically required the officer to deal with his men at the same time as solving concrete problems created by things. Two different kinds of situations were used, one with the candidate in independent command in an outdoor practical situation; the other focussing on his administrative and management roles.

The practical individual situations were designed to bring out certain qualities thought to be related to the capacity to endure stress. They consisted of physical obstacles arranged in a series or “course” with specially constructed apparatus. Athletic prowess was largely irrelevant. The candidate had to assess his own resources in relation to each obstacle. What was looked for was his judgment in overcoming them, as well as his stamina.

munication between military and specialist members, while allowing freer exploration of preconceptions and conflicting beliefs about "officer quality." Differences of opinion were not always resolvable. The difficulty of collating independent reports for the first time at the final conference was one of the main reasons which led to the abandonment of having completely independent roles for the judges.

The experimental board had only just begun to get a program under way when it emerged that 15 to 20 boards would have to be established within the next six months, each with a throughput of 80 to 100 candidates a week. The application of the new methods would not be possible unless the necessary staff could be secured and trained, and the testing program modified to meet such urgent and large-scale needs. To occupy the military roles—president, deputy president and MTOs—suitable officers could be found and trained; but filling the technical roles posed a problem. The psychiatrists' interviews would constitute the main bottleneck. Fortunately, the increasing skill of psychological assistants in making personality pointers enabled the psychiatrist to obtain a sufficient preview of the candidate's personality to distribute his interview time more economically, considerably increase the number of candidates he could see in a day and still feel a reasonable degree of confidence in his judgment.

With a staff of president, deputy president, two psychiatrists, assisted by two (later three) psychological assistants, the interview load could be carried. As regards the practical tests, the MTOs were increased to four. A trial program capable of extension to new boards with intakes of 40 was organized as follows: the psychological tests were confined to the first (half) day; on the next two days the president and deputy president each interviewed 20 candidates, being provided with the intelligence results and notes on the biographical questionnaire; the two psychiatrists each interviewed 20 candidates with the aid of the pointers. All four interviewers prepared reports by the end of the second day for the final board meeting next morning. The MTOs' tests proceeded as before, all candidates carrying out individually prescribed physical tests and taking part in both indoor and outdoor command situations.

This program was accepted by the army authorities and the decision was taken in April 1942 to convert existing CIBs to the new style and to add a number of new boards.

### *Operational Development and Expansion*

The boards were brought under direct War Office control in order that officer selection could develop as a centrally co-ordinated activity. They were named War Office Selection Boards (OCTUs)—WOSBs for short. Simultaneously, the composition of the boards was altered. The psychiatrists and the MTOs

were not members of the board but acted as advisers to a board which retained the composition of CIBs. Visiting members were dispensed with. Nonetheless, their presence had been valuable. They saw what was being attempted and took back to their units first-hand impressions of the work being done.

The changeover from old to new took place rapidly, staff being trained at the experimental board (renamed No. 1 WOSB) and at one of the earliest of the new boards to be established. There were 15 new boards at work by September 1942. There were not enough psychiatrists to have two per board. Commissioned psychologists were appointed, who interviewed the less problematic candidates so that only one psychiatrist became necessary.

The staff of the experimental board had a continuing concern with the inadequate nature of the tests used by the MTOs and with the feelings of the president and psychiatrist that their judgment would be improved if they themselves could see something of the candidates in action. Conversely, the MTOs needed to know something of the inner man so that their cross-sectional view could be better interpreted. The unsatisfactory system of bringing independent reports together for the first time at the final conference required revision.

#### THE LEADERLESS GROUP METHOD

The general difficulty with the MTOs' tests was that they had little or no coherent conceptual framework governing their content and sequence. The method of leaderless groups in which a group was left to its own devices in coping with a situation with which the MTO had confronted it, or which it set for itself, was conceived by W.R. Bion (Bion, 1946; Trist, 1985). Formal leadership was removed and leadership patterns were left to emerge through a series of group situations, beginning with the least structured and proceeding to more structured events. The aim of the leaderless group tests was to reproduce those aspects of an officer's job principally concerned with his approach to, and his relations with, others.

While other methods and interviews informed the testing officers to some extent about the quality of the candidate's social relations, the leaderless group method forced the candidate to reveal this quality directly in the here-and-now. The method made use of the candidate's anxiety to do well for himself, to further his own hopes and aspirations. In individual tests his desire to do better than other candidates presented no problem, but when he was put through tests as a member of a group without a leader, a problem was introduced. The anxiety to look after his own interests remained, but the MTO's instruction called into activity not individuals, but a group formed by those individuals. Moreover, no indication was given as to whether judgment would be on the

## PSYCHOLOGICAL TESTS

The best arrangement was for the psychiatrist to provide an integrated technical report that combined clinical assessment with the more objective measures which psychological tests could give. Four types of test were chosen: questionnaires, intelligence tests, projective tests of personality and individual tests of a laboratory type based on those used in the German army (Ansbacher, 1941). The latter were subsequently dropped as being redundant or impractical. The first two were group-administered written tests and the third soon became so. The function of the two questionnaires was to have recorded, for the convenience of both interviewers, the main features of the candidate's scholastic, occupational and military history and, for the psychiatrist and psychologist, more personal information about family history and health.

In choosing and developing intelligence tests, the following factors had to be taken into account: the capacity to reason with both verbal and non-verbal material (the influence of different educational opportunities being reduced to a minimum); flexibility—the test items were arranged so that each problem would be approached afresh; the maximum discrimination should occur among the top 30 percent of the army population. Candidates were separated into those clearly acceptable, those of borderline acceptability and those unacceptable. Individual confirmatory tests (Semeonoff and Trist, 1958) were given to three categories of candidate: those whose educational or occupational record was out of keeping with the test results; those showing unusual discrepancies between performances on the three tests; and candidates of borderline ability who did well in interview or on the MTO tests.

The projective method creates conditions for the total personality—its conscious and unconscious forces and their organization—to reveal itself in a spontaneous way. Certain projective tests were identified for further work because of their ease of group application and assessment. Three were eventually chosen: a modified Word Association Test, a short series of Thematic Apperception Test pictures and a written Self-Description. These were all given in group form after the questionnaires and intelligence tests.

The purpose of the Word Association Test was to explore spontaneous attitudes towards the officer's job. Words were chosen because of the likelihood of their being linked with such attitudes, including the anxieties aroused (Sutherland and Fitzpatrick, 1945). The Thematic Apperception Test (Murray et al., 1938) was expected to throw light on unconscious conflicts revolving around officer/men relations, those in authority and those who might be enemies. The Self-Description (Wittkower and Rodger, 1941) illuminated the candidate's insight into his strengths and weaknesses and how he handled hostile or favorable attitudes to himself.

The written responses to the projective tests had to be interpreted in a

clinical manner and thus gave scope for large subjective influences. Nevertheless, it was soon proved that psychological assistants who had a fairly advanced psychological training in their university courses before the war could be trained to interpret the material along psycho-dynamic lines with a reasonable degree of consistency. This was a crucial finding, because the few experienced psychologists available were required for much-needed research and development. Furthermore, the projective tests were not intended to be used as independent measures but to provide leads to personality features requiring clinical assessment in psychiatric interview. Their limited scientific status was signified by calling them personality pointers. The pointers helped to identify early in the program those candidates on whom the psychiatrist's assessment would be particularly valuable. They threw light on assets and liabilities in a way that enabled interview time to be used most effectively. To produce the personality pointers for each candidate from the four hours of written tests took, on average, half an hour.

#### ADAPTING TO INCREASED DEMANDS

Candidates were given the status of cadets and shared a mess with the MTOs. There was little difficulty throughout the three days in maintaining an informal atmosphere consistent with the basic features of army discipline and custom. Badges of rank were replaced by identification numbers on arm-bands which, while convenient to board staff, indicated to all that judgments would be made on what the candidate was, not who he was.

The president with the two local COs formed the board proper, while the psychiatrist, psychologist and MTO played the part of expert advisers. The president interviewed all the candidates himself and the two visiting officers usually conducted a joint interview. The two psychiatrists each interviewed half the candidates, and the MTO and the psychologist used the tests described. At the final board meeting the independent interview judgments were placed alongside the data from the specialist advisers and a new composite judgment was made.

While this form of board was appropriate for working out test methods, it was not optimum for operational use. Accordingly, a deputy president and an extra MTO were added and the staff soon divided into two boards, each with its own team of president, MTO and psychiatrist. The psychologist and two psychological assistants worked with both psychiatrists. Each team could handle 16 candidates in three days so that a board with two teams could see 64 men a week.

A system of reporting, analytical only to the point of separating facts, interpretations and gradings, emerged. Though not ideal, it facilitated com-

performance of individuals or the performance of the group. The conflict for each individual candidate was that he could demonstrate his abilities only through the medium of others. This being true of everyone in the group, a common purpose was created, namely, to act towards one another so that each would have an opportunity to display himself.

There were two problems set by the leaderless group method—the real or social problem, i.e., to reconcile group purpose with individual aspirations, and the quasi-real or presented problem. Candidates direct their attention to the quasi-real problem which conceals the real problem, so that the latter is only vaguely sensed. The more that candidates accepted the quasi-real problem, the more could the MTO identify what was spontaneous in their behavior and through this get an indication of their cohesive and destructive tendencies. It was not the artificial test, but the real-life situation that the observers had to watch—the way in which a man's capacity for personal relationships stand up under the strain of his own and other men's fear of failure and desire for personal success (Bion, 1946).

The leaderless group tasks were set in a series intended to parallel phases in the formation of a group faced with a common task. These phases would overlap, but to separate them made observation easier. Groups of eight candidates were found to be best, though groups of 10 were manageable. The interplay of personalities was freer and more illuminating when the groups could be made as homogeneous as possible in regard to age, rank, arm, and length of service. A basic series of tests lasting about two-and-a-half hours was evolved to represent four phases:

*Exploration.* The phase of preliminary contact in which members of the group sized each other up and began to know each other, represented in mutual introductions in which each candidate took about three minutes to give personal particulars of himself. This led on to a free group discussion in which the group had to choose a subject which would make for a good argument and then discuss it, 30 minutes being allowed in all.

*Competition.* In this phase the group members were competing for dominance and the group got some experience of its members in leadership roles. This was represented in spontaneous situations in which several military problems were presented in quick succession, the MTO using the immediate outdoor surroundings as material. The situations were not complex enough to call for action by the group as a whole. It was up to each individual to declare his preferred method of participation (or non-participation).

*Co-operation.* Each individual had to learn that only by pooling of resources and setting aside self-centered attitudes and motives could a goal be reached. This was represented in the progressive group task which con-

sisted of a practical problem. Characteristically the group had to carry a heavy and awkward load over a series of obstacles of increasing difficulty, in a military setting with an air of urgency. The group had to cooperate to produce an acceptable plan and build an organization around the most effective leadership it could produce. The group had to sustain its activity over a period of time—30 to 45 minutes—to reach an objective.

*Discipline.* The individual had to identify himself with the group's decisions and subordinate himself to a pattern of organization in which he had to accept the role assigned to him. This was represented in the group game, usually, between two groups, carrying a heavy object, competing against each other for 20 to 30 minutes around an obstacle course.

#### THE OBSERVER TEAM

The leaderless group tests dealt with the general qualities of social relations which concerned all members of the board. There was an advantage, therefore, in their being observed by the full team of selectors—the president or deputy president, psychiatrist and psychologist as well as the MTO. When a group of observers watched the basic series it was almost impossible for them not to discuss what they had seen and thus difficult to maintain strict independence. With shared observations, differences in opinion were aired early on and this was formalized in a “query conference” at the end of the basic series. Each judge noted those candidates on whom his specialist attention would need to be centered.

“The leaderless group method changed the entire character of the WOSB. The board became a learning community which improved collective capacity through the sharing of common here-and-now experiences of the candidates instead of conducting acrimonious and unresolvable debates on independently based judgments” (Trist, 1985). The creation of the observer team enhanced the value of the special contribution which each member would make in the final board conference. It greatly improved the basis for the collaboration of the three types of judge, both in their feelings about each other's role and in their common task. The shared observations indicated to each member his prejudices and biases. The comparisons of judgments on the same data did much to keep standards similar amongst observers in the same team.

With candidates who failed there was a strong desire among board members to advise them and to secure conditions in which they could develop or use their assets to the best advantage of the army and themselves. Thus letters to COs would explain the board's opinion and ask them to give these candidates