RECENT ADVANCES IN SCREENING THE EMOTIONALLY MALADJUSTED*

JOSEPH ZUBIN

New York State Psychiatric Institute, Columbia University
Towards the end of World War I (July 1918) Pershing cabled the Chief of Staff: "Prevalence of mental disorders in replacement troops recently received suggests urgent importance of intensive efforts in eliminating the mentally unfit from organizations new draft prior to departure from United States" (10). Psychologists had not been unaware of the problem and early in the war (April 25, 1917) Woodworth had been appointed by the American Psychological Association as chairman of a committee on Emotional Fitness for Warfare. Woodworth and Poffenberger worked assiduously on this problem at Columbia and after trying out various tests "hit upon the idea of assembling minor neurotic symptoms, as found by psychiatrists in the case histories of individuals who later developed neuroses or psychoses, and tallying up the score of positive answers"(18). The impetus probably afforded by Pershing's cable led to immediate tryouts of the test which eventually became known as the Woodworth Personal Data Sheet, the first screening device to be used by our military forces. The Personal Data Sheet "was intended as a screening device with primary use of the quantitative score, but also with attention to certain 'starred questions' which the psychiatrists in the conference believed would be of significance quite apart from the total score"(18). The Personal Data Sheet went through two major revisions before it was finally authorized by the Surgeon General for a really intensive tryout in the Army, and by that time, about October 1, 1918, the war was nearly over.

In the postwar period the untried method was used widely in schools and industry with inconclusive results. The advent of World War II found a group of men already engaged in experimenting with screening inventories and soon after Pearl Harbor wide use of such tests was made in the navy, merchant
marine, and to a lesser extent in the army. While it would perhaps not be easy to demonstrate that screening shortened the war and reduced casualties, it would be difficult to avoid such a conclusion. To be sure, screening did not prevent mental breakdowns from occurring in the ranks, both here and abroad, but the number of such occurrences, though large and in some instances appalling, would probably have been even more staggering if the initial screening had been omitted.

The success of these devices was so phenomenal that some of the older psychologists who lived through the deflation of the psychoneurotic inventories during the 30's found it difficult to believe that these tests could be 85 per cent efficient. Certainly, the tests were not perfect, many false positives being unnecessarily caught in the screening net and some false negatives slipping by. Nevertheless the screening technique was a success. Although the final count is not complete, it seems reasonably safe to conclude that screening tests, under favorable conditions, will detect the great majority of the patently maladjusted. They will not, however, perform miracles. For example, they will fail to predict the development of a schizophrenic psychosis before any traces of the disease onset are apparent; nor will they predict the amount of stress a given soldier, normal at induction, can withstand before reaching his breaking point. Indeed, the problem of predicting the breaking point under stress has become the number one problem of screening of the future, and thus far the experimentation with miniature stress situations has not been too successful. A very readable description of this problem, as observed during World War II, is given by Brigadier General Cooke in a recent volume entitled "All But Me and Thee"(2).

It is difficult to fathom why the psychoneurotic inventories worked so well during the war but not at all during the peace. Several factors have probably been responsible. First, we developed better instruments. Second, the general population became better acquainted with such diagnostically meaningful terms as epilepsy, convulsions, mental hospitals and psychiatrists; and mental-hygiene education percolated far enough to permit people to admit headaches, dizzy spells, nausea, vists to a psychiatrist or a mental hygiene clinic, etc., without fear of being labeled insane or mentally deficient. Third, the feebleminded were more readily caught by suitable intelligence screening tests and were thus separated more readily from the emotionally unfit. Fourth, somehow the average recruit seemed to be able to confide more in his examiner than the average peacetime sophomore in his college professor. Fifth, the total score was supplemented with the "stop item" device, which Woodworth's committee had designated as "starred items," to catch in the screening net individuals who, though generally presenting an adjusted picture, had undergone some psychologically notable experience, such as a severe head injury, enuresis, hospitalization for mental illness, incarceration, multiple arrests, etc. Anyone who gave evidence of such psychological deviation was placed in the "suspected" group no matter how normal he appeared otherwise.

But perhaps the most important factor was the lower level of aspiration which these inventories adopted. They were, from their very beginning in World War I, not regarded as personality tests; they were merely sieves
separating the recruits into two groups—those who had to be screened further by a clinician in a short personal interview and those who needed no further screening. This modest claim is due to the trend set by Woodworth and his 1918 committee who stated that “The questionary should be administered by the Psychological Examiners of a camp, being given to groups of recruits along with their group mental test, and being scored by the psychologists and the results, so far as significant for the neuropshiatrists, being then submitted to them”(19). Establishing the normality of about two thirds of the candidates without calling upon the time of the clinicians permitted the very small number of available clinicians to devote their energies to the one third who actually needed attention.

As an example of the efficacy of the screening procedure, the experience of the U. S. Maritime Training Station at Sheephead Bay might be cited(8). Out of every 1,000 men processed, 660 required no further interview after their screening questionnaire was scored. The remaining 340 were given the brief follow-up screening interview because they either had a deviant score or some “stop” item on their record. Of the 340 interviewees, 15 had to be discharged from the service, 75 were found to have temporary or mild emotional conditions which required short-term therapy, while the remaining 250 individuals were false positives. In this way, by seeing 250 false positives per 1,000 admissions, it was possible to detect fully 85 per cent of the totally unfit and select the men who could be reclaimed for the service after brief-term therapy. The 660 men who passed through the screening as normal contained in their number three who were subsequently disenrolled from training for emotional reasons. Of these, one had passed through the screening sieve because he was apparently normal at the time of the screening and did not become mentally ill until later, while two had passed the screening sieve by dissimulating their symptoms. Thus, 18 individuals in every 1,000 were disenrolled on emotional grounds, and 15 of these 18 were discovered during the screening test. The data given above do not include those who eventually broke down during the course of service.

The outstanding advantage of this screening procedure is that the screening load can be adjusted to the capacity of the screening personnel available. This is done by raising or lowering the critical cut-off score, on the basis of which men are selected for interview. A series of controlled experiments have indicated the proportion of the total number of candidates that fall above each cut-off point, as well as the rates for false positives and false negatives that correspond to each cut-off point. This flexibility of the technique permitted each station to adjust its load to any degree of accuracy desired.

Whether the personality inventories born of the war will find use in peace is still a moot question. The emergency nature of screening tests should be borne in mind. Beginning with the Biblical screening test administered by Gideon before his battle with the Midianites and ending with the most recent psychoneurotic inventory, screening tests have come in response to some emergency, only to be discarded when the emergency is past. Will the fate of the present tests be different? The prospects of continued usefulness seem brighter now, especially if suitable allowances are made in the tests for the
transition from war to peace uses. It must be recognized, for example, as Hunt(6) has so well pointed out, that the very items which proved so diagnostic in wartime—dizzy spells, convulsions, headaches, homosexuality—do not have as close a bearing on success in civilian life as they do in military life. Many a successful employee may have some of these symptoms and yet be highly efficient in his civilian occupation. There may be other items which would be just as crucial to peacetime success as those “stop” items were for military success.

Furthermore, although the war emergency is past, we now find ourselves in a new type of emergency. Industry is demanding more efficient methods of selection; school and community clinics require more efficient methods of intake, diagnosis, classification and evaluation. This is not said in disparagement of the methods now in operation. These methods were keyed to the smaller demands of the prewar and war periods. They were never intended to satisfy the large scale demand that has arisen. That the training of more personnel alone is not a sufficient answer can readily be seen by noting how rapidly the demand is outstripping the supply. While the demand increases geometrically with the opening of each clinic, the supply of personnel trickles in arithmetically from the training schools. Group screening techniques have again become a necessity, but a new type of screening is required. The men coming to the clinics have already passed one screen; they have come for help. The task now is to classify the applicants with regard to the types of help required.

The personality inventory in its present form is not suitable for this purpose, since it was devised for differentiating normals from deviants and not for differentiating within the deviant group. For this purpose we need new groups of tests. What should be the nature of these tests? A good starting point would be tests of the word association type, the projective type, expressive movement type, and tests of abstraction of the Shipley Hartford Retreat(12) type, etc. Psychoneurotic inventories of the Minnesota Multiphasic(7) variety can perhaps prove useful. But too often in the past it has been found that this type of test is subject to inexplicable variation from sample to sample. Perhaps a question of semantics is involved here. I recall the difficulty we experienced in eliciting information about enuresis. The formula which seemed to work best was: “Did you ever have weak kidneys?” Whether this formula would work in all regions equally well is very doubtful. An inventory couched in Basic English, as Landis has proposed, might perhaps be the answer.

It may be noted in passing that the material of nearly all of these tests had at one time been used for measuring intelligence or achievement. Inkblots had been used by Binet(1) for measuring intelligence and imagination; pictures of the TAT type had been used by Binet(1) as intelligence test material and by Stern(13) and others as tests of testimony (Aussage); incomplete sentences had been used by Ebbinghaus(3) for measuring intelligence; and handwriting scales had been used by Thorndike(16) for measuring achievement in learning to write. In adapting this material to personality measurement, it had to be freed of the restrictions imposed by the requirements of right and wrong answers. In the process, objec-
tivity of scoring had to be sacrificed. Before these materials can be used for group screening, some rigor and specificity must be reintroduced. But this rigor and specificity will be based on the rich experience which the liberated material has accumulated during the past few decades.

Recent developments in the word association test by Goodenough (4) (who utilizes homonyms) and by Tendler (14) (who has developed certain indexes for distinguishing neurotic from psychotic performance) are good examples of the new type of screening test required. Perhaps multiple choice forms similar to those suggested by Maller (9) and by Thurstone (15) can be readily adapted to group screening.

In the projective field the incomplete sentence test, the Levy Movement Cards (20), the Tomkins-Horn Picture Arrangement (17) test, the Rosenzweig PF test (11) and the Franck Sex Symbolism Test (5) are examples. The chief difference between these tests and their classical prototypes inheres in the fact that the new tests are more controlled and the tasks they set before the subject more specific. Indeed, it is quite likely that the utilization of these more specific tests may help reduce the more imponderable tests, like the Rorschach and TAT, to more manageable proportions psychometrically. As an example of what can be achieved, we can look at the Levy Movement Cards.

The patients are told that they are to be presented with cards showing people in movement. Even with this dead give-away of the purpose of the test, some individuals reject the cards, others see only inanimate objects. Psychometric scoring scales have been devised for these cards which are reasonably objective and which have already proved their diagnostic worth empirically in contrasted groups of normals, neurotics and schizophrenics. It can be readily seen that the other factors of the Rorschach test can be similarly treated in separate tests and similar procedures developed for color, form and shading responses. Another feature that Levy has added to his cards is a phantasy test in which, at the completion of the initial task, the patients are asked to write a short story about the people they had seen in movement. This adaptation of the TAT technique to Rorschach material seems to be a most fortunate approach, since it deals entirely with pure projective material unencumbered with the controls imposed by structured TAT cards.

In the field of expressive movement certain outstanding handwriting characteristics, which can be secured from handwriting specimens obtained in the course of the group testing, may prove to be useful in the preliminary diagnosis.

How are these tests to be used? After the applicant has been given his battery of group tests, he is either sent home to return in a day or so, or, if enough clerical and machine scoring aids are available, his record can be processed immediately and presented within a half hour or so to the senior psychologist for study. Armed with the salient facts of the standard battery, the psychiatrist or psychologist can proceed either with individual or group interviews, group therapy techniques, further tests or other approaches in accordance with the obtained results. It is, of course, a program which cannot be launched immediately. Considerable time and effort will be required for obtaining norms and standardization of the selected battery, but such time would
be well spent since it will eventually reduce the diagnostic load and provide more time for therapy. In this way the mass wartime screening procedures can be transformed to detecting the areas of the individual which require attention.

An additional source of information which might prove of considerable diagnostic value in the case of veterans is the previous record of the patient. If the service data could be made available to each clinic in a master file of IBM cards, much significant material bearing on diagnoses could be obtained. Perhaps a teletype system in central headquarters could feed such information to outlying clinics. Care should be taken, however, not to let the history alone determine the type of therapy and treatment to be given.

Besides the administrative usefulness of such screening procedures in increasing the amount of time available for therapy, other virtues inhere in this approach. It utilizes techniques which the novice can apply, and for a long time to come the bulk of our first interviews will perforce be in the hands of novices. It will permit the trainee to advance, as he gains more skill, from routine administration of group tests to more skill-requiring administration of interviews, individual tests, and finally to evaluation of the total record of a patient and the making of diagnoses and carrying out of recommended therapy.

The standard battery will also provide standard research data on the quality of intake and, if the tests are repeated before discharge, on the particular type of test improvement noted. In this way both the caliber of the therapy as well as of the tests can be improved in a demonstrable, objective manner.

The values involved in such a program cannot be preserved if the standard battery is permitted to crystallize at a given level and to be maintained by sheer inertia from then on. To avoid such a contingency it is necessary to expand the team of psychiatrist-psychologist-social worker by the addition of a research-minded person, be he social worker, psychiatrist or psychologist. His clinical load must be such that he will not be swamped by the regular demands of the clinic, but will have time and opportunity to maintain a running survey of the efficacy of each of the instruments and revise them as need arises. Without such a worker specially delegated to this task, the screening program cannot succeed. The primary task of this research person will be to provide the proper integrating device for encompassing the findings on the various tests. Total scores have long passed their zenith in psychometrics. Pattern analysis, profiles and various other devices will have to be invented, which are simple in their application yet significant and trenchant in their meaning for classification and understanding of the patient.

It might be argued that the clinic patient coming for help will shy away from all of these screening devices and will consider them either as an affront to his dignity or as irrelevant to his problem. Such may be the case in a small proportion; but we must remember the long waiting lines in front of our clinics, the fretful idling away of time in the waiting room for the person who is already worried and anxious. If certain hours could be established as intake hours, at which time all newcomers would be seated and given an orientation talk indicating that this is the beginning of their medical examination,
few if any would refuse to cooperate. There is no one more ready than a neurotic to tell his troubles even to a questionnaire or to become absorbed in a test which may tell something about him. This screening technique might prove even more helpful in medical and surgical clinics, since a large proportion of somatic complaints are usually found associated with mental symptoms.

What degree of success can be expected from such procedures? We cannot hope to be as successful in our diagnostic group screening as we were in wartime screening. The military demands are specified and one can build tests to meet the specifications. Maladjustment in life is more complex and criteria or goals of adjustment more diffuse. Furthermore, the accepted psychological and psychiatric diagnostic categories are not too helpful. Perhaps new goals will have to be established; not diagnostic goals, but therapeutic goals. Tests can perhaps succeed in indicating to what type of therapy a given patient will respond. Then the screening procedures will have attained their ultimate goal, since diagnosis, to be useful, is merely a preliminary screening out of the undesirable therapies and a selection of the desirable one.

In conclusion, screening procedures for segregating the emotional deviant from the normal are now a fait accompli. World War II has provided the crucible for testing their efficiency and they have successfully met the test. The next step in screening is to identify within the maladjusted individual his particular areas of difficulty. This is not a mere academic desire; it is a practical necessity, since without such preliminary diagnostic screening our present method of handling patients is bound to break down, even as selection in the draft and armed services would have broken down in the absence of screening. New techniques are required combining the simpler factors involved in questionnaires, projective techniques, expressive movements, abstraction tests, etc., which can be readily adapted to group use. The provision of such tests will save the time of the patient and professional staff, will permit more time for therapy, will provide the trainee with a series of graded diagnostic tasks suited to his growing ability, and will provide an objective standard record of the status of a patient on admission and at discharge. In order to carry out such diagnostic screening, a new member must be added to the triumvirate of psychologist-psychiatrist-social worker in the mental hygiene clinic—viz., the research-minded clinician.

BIBLIOGRAPHY

of Sciences, New York: XLVI, 1946, 559-570.
17. TOMKINS, S. Tomkins-Horn Picture Arrangement Test. (unpublished)
18. WOODWORTH, R. S. Personal communication, April 10, 1947.