Summer 1961

Situational Tests--A New Attempt at Assessing Police Candidates

James H. Chenoweth

SITUATIONAL TESTS—A NEW ATTEMPT AT ASSESSING POLICE CANDIDATES

JAMES H. CHENOWETH

The author is Chief Deputy U. S. Marshal, District of Alaska, Anchorage. Mr. Chenoweth is a Past President of the Alaska Peace Officers Association and a Past Instructor in Law Enforcement, Anchorage Community College, University of Alaska. His earlier article on eye witness identification recently appeared in this Journal.—EDITOR.

On September 29, 1829, the Metropolitan Police of London, England, entered upon their first day of duty. Of the first 2800 men recruited into that organization, at least 2238 (or approximately 80%) had to be dismissed from the force. All 2800 officers had been hand-picked by a very careful system of selection. Each candidate had to submit three written testimonials of character, one of them being from his last employer; the writers of these testimonials were personally interviewed. If a candidate passed through this stage, he reported for a medical examination, which in practice meant an inquiry into both his physical qualifications and his general intelligence. Less than one in three of the applicants was successful in passing through this stage. Those who did were then interviewed by an experienced personnel officer who eliminated the candidates obviously not suited to police work and passed the survivors on to the first two Commissioners of the Metropolitan Police, who again interviewed the remaining candidates. The disapproval of either Commissioner was sufficient to reject the candidate.

The above technique for the evaluation of police applicants was originated over 130 years ago; it is still the basic examining procedure used by many police agencies today. In it we may glimpse the seeds which were subsequently transformed into our present methodology of assessment; the personal references, the background investigations, the physical and mental tests, the oral interview. Nevertheless, it remains lamentably true that in spite of our advances we have not successfully bridged the gap between the portrait drawn for us by the police applicant and the realistic picture of that same applicant engaged in police activities. Too often we have found that a candidate whose personality, capabilities, and prior experience seem admirably suited to police work is, in reality, ill-equipped to meet the rigors of a demanding profession. Aware of the necessity for eliminating these individuals from police service, many agencies are giving increased consideration to the inclusion of psychological testing techniques in their recruitment and assessment programs.

It may be well at this time to review briefly some of the basic principles of psychological testing. In general, a psychological test is a method for evaluating a person's behavior. The analysis of test results has been defined as a systematic and objective procedure for comparing the differences in behavior between two or more persons. By conducting psychological tests we attempt to predict the future behavior of an individual and to discover within that same individual the differences between his various characteristics. By and large, the police profession is more interested in the predictive aspect than in the diagnostic aspect; we are more concerned with the effective selection of recruits than we are in determining the causative factors that result in rejection.

Psychological tests can be divided into two general categories; those which measure maximum performance and are used to determine an individual's ability and capacity, and those which measure typical performance and are used to determine an individual's habits and personality. Police agencies have a need to know the information obtainable in both of these categories; that is, we need to know not only what a candidate could do (given the right training and guidance), but also what he does do in immediate, typical, day-to-day situations.

During the process of constructing assessment and examination programs, the police profession has—to date—paid more attention to those testing techniques which are designed to reveal maximum performance than to those designed to reveal

typical performance. Tests of maximum perform-
ance include the following: The Stanford-Binet
Intelligence Scale, the Wechsler Adult Intelligence
Scale, the Army Alpha Intelligence Test, The
Civilian Edition of the Army General Classifica-
tion Test, the Henmon-Nelson Tests of Mental
Ability, the Otis Self-Administering Tests of
Mental Ability. These are generally referred to as
intelligence tests.

In his book *Psychology for Law Enforcement
Officers* George Dudycha says, "The selection of
policemen on the basis of intelligence alone is not
enough. An applicant may be superior in this
regard and still be a dismal failure as a police
officer because of other traits he possesses."
Loosely speaking, tests of typical performance (as
distinguished from tests of maximum performance)
are the media used to discover these "other
traits." Tests of typical performance may be
subdivided into three groups: projective tests,
self-report tests, and behavioral observations.

Projective tests make an effort to probe beneath
the systematized behavior patterns each individual
acquires in concealing his true individuality; they
attempt to penetrate this barrier and determine
the underlying personality structure. Examples of
projective tests are the Rorschach Test and the
Thematic Apperception Test. Projective tech-
niques are still being explored, and the practical
uses of this method of testing have not been fully
developed. As Doctor James Rankin indicated in
his article on "Preventive Psychiatry in the Los
Angeles Police Department," (which appeared
in the July-August 1957 issue of *Police* magazine),
"there have been many conflicting opinions about
the validity of Group Rorschach and a number of
different techniques for its use." Apparently one
of the chief disadvantages of this test is that a
highly trained specialist must administer, score,
and interpret the test.

By contrast self-report tests are relatively easy
to administer and score. Unfortunately, they
also depend for their validity upon the personal
insight and veracity of the applicant. Because they
require the candidate to evaluate himself, they
are based upon the unproven assumption that
the applicant is willing and able to both know
and reveal the truth about himself. Probably no
single profession has greater awareness of the
immense void between what a person says and
what a person does than the police profession.
Certainly, the major part of our activity in resolv-
ing police incidents involving people consists of

evaluating what a person has said in relationship
to what he has actually done. The family disturb-
ance, the traffic accident, the suspicious loiterer,
the victim, the witness, the suspect—all of the
people involved will eventually say something
which must be measured against our knowledge
of their actual performance. Naturally enough,
police applicants subjected to self-reporting tests
will also do their utmost to put their best foot
forward; one of our difficulties arises when we
must base our predictions of future performance
upon the immediate but unproven assertions of
the candidate. This flaw is especially pronounced
when the candidate is a psychological sophisticate.

Another objection that has been voiced against
self-report tests is that they become unreliable if
the questions used mean different things to
different people. Many of these tests use words
such as "always", "frequently", "usually", and
"often." The ambiguity of such words casts doubt
on the accuracy of the conclusions drawn from
such tests. In spite of Doctor Rankin's expression
of increasing confidence in tests such as the
Minnesota Multiphasic Personality Inventory,
some psychologists have asserted that self-report
methods should never be used where acceptable
substitutes are available.

One type of acceptable substitute is behavioral
observation, which is an attempt to study the
subject in action. The objective scrutiny of a
supervising officer as a trainee progresses through
his probationary period is an example of behavioral
observation in a "natural" situation. There are
critical objections to be overcome in using this
technique. As Lee J. Cronbach states in his
Essentials of Psychological Testing:

"Observation in normal situations escapes the
errors of self-report, only by introducing marked
observer errors. Field observation is frequently
impractical because of the large amount of
observing required for reliability, and has the
great disadvantage that it is impossible to
compare subjects or traits not normally evi-
denced in their daily activities. If, for example,
an investigator intends to study individual
differences in behavior after long periods of
wakefulness, he can gather little evidence by
field observation of workers or students. Only
by setting up an artificial situation in which
each person is kept awake for a long time can
he observe how the person reacts to such
fatigue."

As distinguished from observations in normal

situations, there are distinct advantages to be realized from observations in test situations. A test situation, or "situational testing," is a technique through which a candidate is exposed to a carefully constructed situation; his reaction to the situation produces behavior that may predict his reaction to comparable situations in the future. The stimulus situation must be as nearly uniform as possible for all candidates. It must be designed to permit variant behavior from different subjects. It must be so constructed that the candidate is unaware of which characteristic is being observed. The method of performance is noted as well as the amount.

In *Essentials of Psychological Testing* Lee Cronbach has this to say about the situational testing technique:

"The greatest advantage of the test observation is that it makes possible the observation of characteristics which appear only infrequently in normal activities—characteristics such as bravery, reaction to frustration, and dishonesty. A single situational test may reveal more about such a trait than weeks of field observation. Second, the subject's desire to make a good impression does not invalidate the test. In fact, just because he is anxious to make a good impression, he reveals more about his personality than would normally appear. It is necessary, however, to take this motivation into account in interpreting results. The third advantage of the situational test is that it comes closer than other techniques to a standardized measure of typical behavior."

Cronbach goes on to say that the principal uses of situational tests have been for research in character, frustration, and that they are especially helpful in studying thinking habits and reaction to emotion-producing situations. It seems obvious that police agencies have a need to acquire this same information about the applicants they process. This need has been formally recognized by police assessment personnel and to some extent every agency has created its own situational tests. For example, oral interview is easily adapted into a stress interview, of which George Dudycha has said in his *Psychology for Law Enforcement Officers*, "The purpose of this method is to place deliberately the applicant in a situation that creates stress of a verbal and motor sort, and then to observe and rate the person on various characteristics." Inspector Robert Gallati of the New York Police Academy noted this same technique in an address he delivered on November 19, 1957, in which he said: "The oral interview by superior officers of the police department has many psychological implications. The candidate is not unlikely, under the stress of such an interview, to reveal otherwise undetected weaknesses. While it is difficult to delineate the attributes which will determine whether or not a particular individual will prove to be a satisfactory policeman, there seems to exist among experienced officers an intuitive ability for such prediction."

Doctor James Rankin's article on preventive psychiatry, which has already been mentioned, includes the following: "Latent neuroses are encountered rather often. These are the individuals with whom one must use his best clinical judgment to estimate how much stress they can take; whether compensatory mechanisms are adequate; and whether the degree of personality integration is sufficient to take the expected stresses of police work. Frequently in this group we do a bit of fence straddling. In other words we will not reject these individuals outright but will allow them to go into the Academy with warning comments. Traits that should be watched for in the Academy and during the probationary period are enumerated, and we may even suggest that the Academy apply certain types of stress during the training period to see if latent weaknesses might be brought out. . . . Thus, instead of arbitrarily saying that a latent neurotic is not satisfactory, we give them a test of functioning under real stress conditions." Inspector Galloti agrees that the stress situations encountered during the Police Academy training period and subsequent probationary experience in the field, have profound psychological implications.

Not every agency is fortunate enough to possess the facilities enabling them to spend thirteen weeks in evaluating the desirability of retaining their police candidates in their organization. Indeed this system does little to counteract the objection that the candidate performs under observation differently than he would otherwise.

Another objection to the procedure suggested by Doctor Rankin is that many agencies find it more difficult to separate a candidate once he has been accepted. By implication the police department at Berkeley, California, recognizes this difficulty when they advise the officer conducting the Personal History investigation that the key point he should keep in mind is "facts are necessary to justify the removal of the candidate

from an eligible list." Further on the same in-
structions state, "It is a painful experience to
separate a policeman during his probationary
period." Inspector Robert Gallati of the New
York Police Academy suggested, in the address
already referred to, that in areas where civil
service personnel boards are responsible for the
recruitment of police personnel, such boards may
not be in agreement with the principles of assess-
ment advocated by police officials.

This then would appear to be the present status
of situational testing: It is a valid approach to
problem of predicting a police candidate's future
performance but is generally applied in a disguised
form during the oral interview or during formal
training and probationary periods. It seems equally
obvious that little consideration has been given
to the possibility of processing a police candidate
through a series of situational tests as part of
his examination procedure.

Before we discuss the administrative problems
involved in such a procedure, it may be timely to
outline some of these tests and the behavioral
reactions they were designed to reveal.

The man who is generally reputed to be the
"father" of situational testing techniques was a
remarkable psychologist in Hitler's Wehrmacht
named Simoneit. He observed officer candidates in
a number of natural situations, recording his
observations, and analyzing them in terms of the
job for which the candidate was being considered.
If this evidence was inconclusive, he would create
artificial situations as nearly as possible like the
situations the candidate would meet in actual
warfare, and then see how the candidate reacted.

One such artifice is described by Lee Cronbach
in the book previously cited.

"A device for studying social behavior in per-
haps its least complex form is a German test for
leader selection, developed before World War II.
A special apparatus is used, consisting of two
pairs of shears, linked by rods so that they must
move in unison. While one shear is opening,
the other is closing. Each subject (candidate)
operates one pair of shears, cutting a series of
increasingly complex patterns from a sheet of
paper. The shears are so arranged that if one
man goes directly and forcefully at his task,
the shears of the other man move in a rhythm
which makes accurate cutting almost impossible.
By means of observation, automatic recording,
and inspection of the product, the tester looks
for evidence of initiative, dominance, and

cooperation which is used with other data in
assessing workers or soldiers."

Our knowledge of situational test techniques
owes much to the Office of Strategic Services;
their assessment program during World War II
depended upon test situations for much of its
effectiveness in selecting the right person for the
right job. The O.S.S. utilized a test called the
"Ball and Spiral" which is similar to the "Shears"
situation described above. In the "Ball and Spiral"
test six candidates were asked to maneuver a
large cone in such a fashion that a ball would
roll up a spiral ramp circling the cone until it
reached a shallow platform at the top of the cone.
The candidates were told that it was a test of
physical coordination as well as group cooperation.
Individual performances were ostensibly graded
by penalizing a candidate who allowed the ball to
drop off the ramp on his side. Group performance
was ostensibly scored on the basis of the time it
took the group to complete the task.

Because of its apparent simplicity, the task was
an extremely frustrating one. The cone had been
carefully constructed so that it was exceedingly
difficult to keep the ball on the spiral ramp. In
trying to keep his own penalties minimal while at
the same time he improved the group score, each
candidate found himself working toward two
mutually incompatible goals. The resulting frus-
tration and irritation revealed many subtle
personality qualities. No less revealing were the
hidden observations during the period immediately
after the test when—with the release in tension—
each candidate reacted in accordance with his
dominant personality drives. Sheepishness, perse-
verence, self-assurance, disgust, disdain, aloofness,
wrath, frustration, disappointment —all of these
emotions found an outlet that was useful to the
observer.

One type of situational test has already been
utilized in a police training program in Anchorage,
Alaska, with surprising results. The O.S.S. called
this test "Red is Blue and Up is Down." The
O.S.S. candidates were handed cards containing
two simple instructions to be obeyed in running an
"obstacle" course; unknown to the candidates,
there were two separate sets of instructions, each
set the exact opposite of the other. One set told
the candidate to go to the left when he reached
any obstacle marked with a red sign, and to go to
the right if it was a blue sign; he was to go over
any obstacle marked with an "A", and under
any marked with a "B".

The next candidate in line received a set of instructions just reversing the set already mentioned; in other words he was to go to the right at a red sign and left at a blue, and under obstacles marked "A" and over those marked "B".

The first man to run the course was an O.S.S. "plant" who followed first one instruction, then its opposite. Occasionally, he took paths and turnings not controlled by signs. His object, of course, was to set an example that would be followed by those candidates who would ignore their specific instructions in order to emulate the actions of other candidates preceding them over the course. The exact route taken by each candidate was mapped by observers; frequently candidates were genuinely shocked when they realized how completely they had ignored their instructions. Although the O.S.S. staff were never able to define the combination of qualities producing success or failure on this test, they believed it was of great value to know how a man reacted to this test before a decision was made sending him into the field.

Not too long ago police officers attending a training class in Alaska were told that they were about to see a prepared skit that was being presented to illustrate their fallibility as eye-witnesses. They were advised to take nothing for granted "from this moment on", to view everything with suspicion and skepticism for an effort would be made to deceive them. Then they were told that in order to resolve future debate as to what had actually occurred, a motion picture camera would record the short skit. After the skit was enacted, the students filled out questionnaires as to what they had seen.

A week later the movie was shown to the students, after which they were requested to answer truthfully and from their own memory additional questions as to what had actually taken place during the skit. What the students did not know was that the movie was a fake, having been carefully prepared in advance to include many elements in marked contradiction to the events of the actual skit. A subsequent comparison of the two sets of questionnaires revealed that an overwhelming majority of the students accepted without question the film version even when it was in obvious opposition to their own memories of the skit.

Sometimes the purpose of each situational test was hidden more effectively if it was removed from the classroom. An ingenious example of

this was the "Brook" test. Using a narrow stream with trees on both banks, a rock, a log, boards, ropes, a pulley, and a barrel in the immediate area, candidates were told that the stream was a raging torrent, the log was a delicate range-finder, the rock was a box of percussion caps. Using materials in the area the team of candidates was to transport the rock and log across the stream. Easy solutions of the problem were ruled out by a fanciful elaboration of the physical situation. Watching the group in action on this problem, staff members of the O.S.S. were able to rate certain variables of personality such as energy, initiative, effective intelligence, social relationships, leadership, and physical abilities. Needless to say infinite variations under these general headings were observed.

Again mixing fancy with fact, O.S.S. personnel devised a test called the "Wall" in which candidates were faced with the problem of climbing a wall ten feet high and, by using a log, boards, and ropes in the area, travel from the top of the wall to the top of a similar wall, eight feet away and parallel to the first wall, without touching the ground between the two walls. This test immediately followed the "Brook" test and was designed to reveal the same variables. The O.S.S. staff treated the two tests as a unit, their final rating being based upon a candidate's performance in both situations.

Perhaps the most elaborate of the situational tests used by the O.S.S. was the "Construction" test. In this one, each candidate was required to construct a simple frame structure out of a collection of wooden materials that resembled a large Tinker-Toy set; the hitch was that he could not do the work himself but must direct and supervise two helpers, who appeared to be simple laborers working nearby. In reality these two helpers were members of the O.S.S. staff. Although the helpers had to carry out any explicit order given them by the candidate, their main function was to present him with as many obstructions and annoyances as possible within the required period of time. Apparently, they were so successful at this that no candidate ever completed his task in the allotted time.

There was, of course, no "book solution" for the candidate under these circumstances. Apparently a test enabling the candidate to demonstrate his leadership, it was in reality a test of his capacity to work toward a completion of his project in spite of the increasing pressures supplied by his

"helpers." The presence of observers placed the candidate under additional strain. It is interesting to note that although a number of candidates guessed the true nature of the test and the role of the "helpers", the test generally revealed some insight into their personalities too.

The tests outlined are merely representative of the type and variety of situational tests, and do not by any means exhaust the ingenious situations contrived by the O.S.S. and other testing groups. But it is time to consider the principles which make these tests a valid media for predicting future behavior. One of the most fundamental of scientific assumptions is the principle of consistency which states, according to *Assessment of Men* which was written by the O.S.S. Assessment Staff, that the "interactions that occur in two identical situations will be identical, or more specifically, that a given subject will respond to similar environmental situations in a similar manner." While admitting that it is not possible to obtain the scientific ideal in these matters, it seems that we can "expose a man to a variety of situations of the same type as those he will meet in the field and, allowing for certain expected developments in his personality during the coming months, predict future performance level on the assumption of consistency." It is not intended to imply that situational testing techniques supplied the sole criteria for the O.S.S. assessment program. Quoting the previously cited *Assessment of Men* again, we find the following: "For a short over-all assessment the interview is probably the best and only indispensable method we have, but many others are very useful: Intelligence tests, projective tests, questionnaires, autobiography, informal observations of behavior, and situational tests." The purpose of this paper has been to point out benefits obtainable through the utilization of situational testing techniques in the primary stages of recruit evaluation, and not to suggest the elimination of other proven techniques.

The task of evaluating police applicants would be considerably simplified if we could itemize and isolate those specific qualities and characteristics that make a good policeman. Most critics are ready to list those qualities that make a bad policeman, such as dishonesty, insecurity, etc. Frankly, we do our profession a disservice by this negative approach for some qualities are most obviously necessary ingredients in the make-up of a good police officer. May we list the following:

Energy and initiative, effective intelligence, emotional stability, social relationships, leadership, security, physical ability, observing and relating, and certain propaganda skills. If these phrases sound familiar, it is because the O.S.S. assessment technique was designed to reveal the presence or absence of just these qualities. It follows then that adapting some of the O.S.S. situational tests for use in evaluating police applicants would be a worthy and effective procedure. Here at least is a stepping-stone that might point the way toward the creation of situational testing procedures designed specifically for assessment in police work.

Obviously, the feasibility of such a project must rely, to a certain extent, upon the success of the O.S.S. assessment program itself. According to Wm. J. Morgan's book *The O.S.S. and I*, candidates who had successfully passed the O.S.S. assessment program at Pemberley, England, were frequently passed on to other phases of advanced training. Prior to the O.S.S. program five out of ten men failed the advanced training; after the creation of the O.S.S. program only one out of ten failed. But systematic studies on the reliability of situational tests are few and comparatively inconclusive. On this subject Lee Cronbach has the following to say: "At this time, it is impossible to make a general evaluation of the validity of situational tests. So long as they are treated only as objective measures of limited traits, few questions arise. Evidence is quite inadequate to support any contention for or against their validity as measures of the total personality. As predictors, they seem to have promise, according to military experience." And further on, he makes what might well be the most objective judgment available at this time: "We must necessarily await further research before generalizing about observations in test situations. Situational and projective tests may be the only truly valid testing approach to personality." Progressive police organizations could augment the necessary research by the deliberate inclusion of situational testing in their assessment programs.

What are some of the factors affecting the feasibility of including situational tests in the examination stage of recruitment?

First of all, expense. Most police agencies operate with their backs figuratively pushed to the fiscal wall. Can they justify the additional expense of situational tests? But how much

expense is really involved. As indicated by the tests already cited, a great variety are available, ranging from simple elaborations of the oral interview to the elaborately structured tests utilizing expensive materials and facilities. Tests such as the "Shears" and the "Ball and Spiral" need only simple devices that could be fashioned in any home workshop. Any ordinary room would provide ample space for conducting many of these tests, and most police agencies have an observation facility whereby post-test observations could be made without the candidates' knowledge. Of course, police organizations with access to training academies could easily obtain the space and equipment necessary for the more elaborate tests.

Second, specialized personnel to administer and evaluate the tests. It is certainly true that the reliability of the data obtained from these tests increases with the training and experience of testing personnel. But what are the present alternatives? Projective tests require the services of highly trained personnel, not only to administer but to score and interpret. Self-report devices depend upon the reliability of the candidate. Observation in normal situations relies upon the untrained observer and takes place under normal, everyday circumstances; such situations provide little indication as to a police candidate's behavior under the stress of the unusual and infrequent incident which may result in tragedy or social disapproval. These normal situations seldom reveal any symptom of dishonesty, cowardice, over-aggression, insecurity, or many other traits which are detrimental to the fulfillment of police responsibilities. Situational testing would create the unusual incident in a disguised form, revealing behavioral patterns and responses that need not be interpreted by specialized testing personnel but are recognizable by any trained observer. Anger, frustration, contempt, excitability—these facets of the personality are relatively easy to recognize. Anyone who is in a position to evaluate an officer's behavior in carrying out his official duties should also be able to detect that same behavior if it occurs in a test situation. The reasonable approach would seem to be to select from the personnel of each agency those officers who have demonstrated some talent in the area of behavioral observation; then provide them with sufficient additional training to enable them to extract the maximum behavioral data obtainable through situational testing techniques.

It seems strange that in our present examination of police candidates, we place emphasis upon the oral interview which is frequently conducted on a spontaneous, intuitive basis by officers who have had no training in the field of stress interviews, and upon background investigations which depend for their accuracy upon the many persons with whom the candidate has had some prior contact but who are rarely qualified to give objective and unbiased comments concerning the candidate's past conduct. Without demeaning the information obtainable through both sources, one cannot help but believe that situational testing would be a meritorious addition even if it were administered by untrained personnel.

Third, test construction. It is not necessary here to list the infinite talents that are essential ingredients in every good police officer. Most of the situational tests devised by the military provide the type of behavioral data needed in police work. Thus we have at hand a wealth of information and techniques that, with minor alterations, can be put to immediate use. These tests should be administered to three different groups at first: The good police officers, the poor police officers, and "non-police" civilians. We should encourage research institutions, universities, and graduate students of police science, police administration, and psychology to undertake the task of standardizing and validating the results of such tests with the ultimate aim of devising new situational tests directed specifically at the requirements of police work so that ultimately we would possess a repository of test situations and techniques which may be utilized by large police agencies or small, rural or urban, isolated or integrated with other agencies. This approach would truly justify our claim to a professional status.

One final word. In an effort to present the possibilities inherent in situational tests, the information and suggestions have been presented in the simplest possible manner. If they arouse any interest at all, the reader is urged to go directly to the two books upon which the writer, as a layman, relied heavily for my information, *Essentials of Psychological Testing* by Lee Cronbach, and *Assessment of Men* by the O.S.S. assessment staff. The wealth of information contained in these two works should provide stimulation and encouragement to police administrators desirous of improving their assessment procedures.