

## **Differential Prediction of Success through the use of Psychological Tests**

**S. Chatterji & M. Mukerjee**

**Indian Statistical Institute, Calcutta-35.**

Examination of the abilities required for success in various fields of endeavour shows that each measurable aptitude is usable in a number of occupations and hence it is economical to construct standard test batteries from which it is possible to get scores indicating an individual's expected success in a number of occupations by utilising the test scores. There are several methods of utilising the test scores for classifying the individual and these methods are briefly discussed here. From the literature it is further observed that most of the aptitude tests developed in India are in a developing stage and even with respect to other tests of foreign origin not much systematic work has been done as far as differential prediction is concerned. Finally, the predictive validity of the Aptitude Test Battery for Differential Prediction which is developed here is briefly discussed.

In a vocational or educational guidance situation it is not sufficient to indicate that an individual possesses certain amount of Verbal ability or Mechanical aptitude or Spatial sense. The counsellor has to find out how far the candidate is competent in terms of his abilities, to succeed in specific subjects or fields of study or occupations. For this purpose the pattern of the existing abilities as present in the individual, is important because success in any educational or vocational field depends upon the specific combination of different abilities.

Examination of the abilities required

for success in various fields of endeavour shows that each measurable aptitude is usable in a number of occupations and hence it is economical to construct standard test batteries from which it is possible to get scores indicating an individual's expected success for a number of specific occupations by suitably utilising the test scores.

Differential prediction is the evaluation of an individual's relative chance of success in two or more fields of endeavour with the help of a common battery of tests. In the academic career of a student, there comes a time when he has to think "for what occu-

pation or educational field am I best fitted?" as is very aptly summarised by Bingham (1937). Hence, it is but natural that the parents as well as the young people themselves want to know, as accurately as possible, the requirements in the various professional trainings and the clues towards probable achievement in field or fields of ambition.

This problem of fitting the right type of person to the right type of vocational or educational field is also considered from another point of view by several educational institutions. Different institutions have to face the problem of choosing, from among the applicants, a group who as per all expectations shows the greatest promise. For this reason they restrict admission to those who do not seem to show signs of success to an adequate degree in the field concerned. To determine this, special aptitude tests as well as tests measuring previous educational achievement are generally used.

Personnel selection and educational guidance usually go hand in hand. The process of ascertaining a candidate's fitness to undertake a course of professional training is in practice the same whether it is for either selection purposes or for providing individual guidance as is stated by super (1962).

"In either instance it is desirable as a first step to appraise the individual's records of accomplishment in the subjects he has studied and his performance in

tests of scholastic aptitudes and from these data to estimate the probability that he is capable of mastering courses of study in professional levels".

Thus, the trend in aptitude testing so far has been directed towards the measurement of specific abilities by a battery of tests and the judgement is based on the obtained pattern. The underlying principal is, as stated earlier, that each measurable aptitude can be used in predicting the success in a number of different fields. Hence, such tests will not only help the student in detecting whether he (or she) is suitable for a particular course, but will also help in deciding which course is least suitable for him and what are the probabilities of his success in different fields. Again these tests can also be used for obtaining maximum efficiency in situations where each individual may be eligible for several assignments.

For the best utilisation of the information obtained through administering a battery of tests, several statistical methods are available and these methods are briefly discussed here.

#### **Methods used in Differential Prediction :**

After obtaining the scores of the candidates on different tests included in the battery, the counsellor faces the problem of interpreting the scores. Though the subjective judgement given by an expert counsellor is proved to be correct in many instances still it is

always better to have an objective check in this respect.

Of the several methods for classifying the individual into a suitable group on the basis of his scores on the tests included in the battery the more important and popular ones are as follows : (A) Regression Analysis (B) Discriminatory Analysis and (C) Profile Matching. A brief description of these methods is presented below.

(A) *Regression Analysis* : This method requires data from several groups, where each group undertakes one particular educational or vocational training course. The battery of tests in question is administered on all these groups of candidates and separate regression equations are developed for each of the groups separately on the basis of the available measure of success in that particular field for them. Now when a particular candidate comes for advice, this battery of tests is administered on him and by using the regression coefficients separate composite scores are obtained for separate field on the basis of the test scores. Each of these composite scores is the expected score of the candidate indicating success in that field. Unfortunately these composite scores are not directly comparable as the distributions of the criterion scores (measure of success) in different fields may vary. So the counsellor faces the problem in deciding which one of the composite scores indicates the best performance on part of the candidate. This problem, however, can be solved by

getting the distribution of composite scores for each group and dividing it into several parts on the basis of the standard deviation. Letter grades can be given to these parts which are directly comparable.

(B) *Discriminatory Analysis (Identification)* : This method is applicable when the problem is to assign an individual to one of a finite number of groups on the basis of a set of measurements of the individual in question. Suppose there are  $p$  tests included in the battery. The problem is simplified under the assumption that the joint distribution of the scores on the  $p$  tests is a  $p$ -variate normal in each of the groups in one of which the individual is to be classified. For a satisfactory solution of the problem we need to know the following :

- (i) The mean values of all the tests used in the battery for each of the groups.
- (ii) The variance-covariance matrix among the test scores.
- (iii) The proportions of the individuals belonging to these groups in the common population to which the candidate in question belongs.

When all the above information is available for each group a linear discriminant score 'S' (which uses all the statistics mentioned above) can be obtained and then the individual is assigned to that group for which this discriminant score is highest. It can be proved that when the decision is based on the discriminant score, than frequency of wrong

identification is at a minimum.

When there are only two populations, then only one comparison is involved and here the decision can be taken by computing the difference score  $S_1 - S_2$ . This gives the linear discriminant function as derived by Fisher. Rao (1965) suggests that individual discriminant score should be computed even when the problem is to assign the individual in one of the two groups since they are more informative.

It should be noted here that some difficulty crops up when the battery consists of aptitude tests only. Though it is claimed that aptitudes are independent of training or experience, yet it is observed that the tests generally included in the aptitude test battery measure something which may change to some extent due to training and experience. Bennett's study of the effect of previous training in Physics upon the scores on his Mechanical Comprehension test can be quoted in this respect. In this he states that, "315 applicants for defence-industry answered a question concerning previous training in Physics. The 220 persons who had such training made a mean score of 41.7 while the 95 reporting no training made a mean score of 39.7" (cf. Buros, 1959).

Naturally it becomes a problem to get a group mean of the aptitude tests. The question is whether the mean should be obtained on the basis of the freshmen or on the basis of the scores earned by those who have com-

pleted the course and hence do not belong to the group. But at the same time it can be noted that the scores made by the members of the group after completing the course will be different from what they would have scored without the training.

The problem of getting group mean can be solved in the following way. In case of educational training, the scores of those admitted in the course can be obtained and they are followed throughout the training period. Only the scores for those persons who successfully complete the course may be taken into account while computing the group mean. In this way one can eliminate the training effect on the score if such effect exists. Here, the only difference is that instead of administering the tests on the successful candidates at the end of training the test is administered on the same group at the beginning of the course.

Now the proportions of persons in different groups into which the individuals are to be classified in the common population can be obtained by getting the numbers of persons successfully completing different courses among those who are coming for guidance. For example, the proportion of students (successful) in the Science stream of the Higher Secondary Examination can be obtained if the number of students successfully passing the Higher Secondary Examination from among the students reading in Class VIII can be obtained, if it is desired to provide guidance to the students at the Class VIII

level. These numbers can be obtained by just following the students for four years after they are promoted to class VIII.

(C) *Profile Matching*: For applying this method the average scores on different tests included in the battery obtained by different professional or educational groups are to be known. When these averages are plotted against each tests and these points are joined we get the typical profile for that group. Now the individual seeking guidance, can be put in that group with which his pattern of scores matches most. The index of pattern similarity can be obtained by using any one of the several methods suggested in this connection as it has been empirically proved that all these methods are equally good (Helmstadter, 1957).

As stated in connection with Discriminatory Analysis the averages of the scores on the tests should not be obtained by administering the test at the end of the course. Instead these tests should be administered at the beginning of the course, and when they have completed the course the scores of those who are successful are only taken into account for obtaining the group mean.

#### **Aptitude Tests Batteries available for Differential Prediction :**

A brief description of several representative tests used for differential prediction and relevant information regarding their validity is presented here.

(1) *Differential Aptitude Tests Battery* (Psychological Corporation) : It is one of the most popular multiple factor test batteries. It consists of seven tests—(a) Verbal Reasoning, (b) Numerical Ability, (c) Abstract Reasoning, (d) Space Relation, (e) Mechanical Reasoning, (f) Clerical Speed & Accuracy and (h) Language Usage. The score obtained with the battery is highly related with the grades obtained as College freshman. Out of the seven tests used in the battery the ones of Verbal Reasoning, Numerical Ability, and sentence part of the Language Usage serve as good predictors of grades obtained in English Mathematics, Science & Social Studies. The Clerical Speed and Accuracy test has good predictive ability with respect to courses like short-hand and typing. For courses like book-keeping, business, arithmetic, physical education, health, home economics and music the sentence test often shows high predictive value. However, Buros (1959) states "Despite the extensive validity coefficients for separate DAT sub-tests, the differential validity of the tests in predicting various criteria is still without substantiation".

(2) *General Aptitude Test Battery* : It contains 12 tests and factor analysis has revealed that there are nine factors operating in the 12 tests. These factors are as follows (G) Intelligence, (V) Verbal Aptitude, (N) Numerical Aptitude, (S) Spatial Aptitude, (P) Form Perception, (Q) Clerical Perception, (K) Motor Coordination, (F) Finger Dexterity

and (M) Manual Dexterity. Validity studies are being continually carried out for improving this battery. Several research studies have been conducted by various investigators and according to Buros (1959). "On the basis of the statistical findings and other considerations, critical aptitudes are selected for the occupations and cutting scores determined for each aptitude. This is done in such a way as to eliminate about one-third of the individuals. An attempt is made to fit the occupations into already established occupational-aptitude patterns". The statistical methods adopted in the validity study or in setting up the cutting points are usually criticized by many reviewer of the battery. They point out that carefully designed factorial studies based on a large amount of data are needed to verify the factorial structure of the test battery.

(3) *SRA Primary Mental Abilities* : This battery contains tests applicable at three different levels classified according to age and school grade. At each level the battery is capable of yielding scores on five dimensions. It has been observed from the validity study that the V and R scores are most reliable and have the highest validity coefficients. With the ordinary school examination marks or college grades the test scores are not so highly related.

(4) *Holtzinger-Crowder Unifactor tests* : This battery is composed of nine tests which can be grouped into five categories according to the dimension or factors measured by them. These categories are verbal spatial, numerical

reasoning, and scholastic aptitude.

In order to study the validity of the several factor scores they were correlated with high school grades, ratings given by teachers and with the performance on several available standardized tests. Most of them are concurrent validities and a few of them measure predictive validity. On the basis of the obtained validity coefficients, different formulas with different sets of weights have been developed for predicting success, e.g., achievement in science, achievement in social studies, achievement in English, achievement in Mathematics etc. This battery can be profitably used for educational guidance. A score for predicting general scholastic aptitude can be derived from the factor scores and corresponding I. Q's can be estimated.

(5) *Yale Educational Aptitude Test battery* : This battery has been mainly designed for aiding educational guidance in schools and colleges. It includes seven tests and is suitable for grades 9 to 16. Differential validity of each part has been checked mainly against grades in the appropriate college groups and these coefficients are generally high. But while correlating with appropriate courses the coefficients are generally under .20. Available data on validity indicate unusual promise for differential prediction of academic performance in broad curricular areas.

(6) *Gullford-Zimmermann Aptitude Survey* : This battery is composed of seven tests (i) Verbal Comprehension, (ii) General Reasoning,

(iii) Numerical Operations, (IV) Perceptual Speed, (V) Spatial Orientation, (VI) Spatial Visualisation and (VII) Mechanical Knowledge. Results from the factor analysis performed indicate that the tests are more or less factorially independent. But nothing could be said about the predictive value of this battery as the validity study has been conducted on a very small sample.

(7) *Flanagan Aptitude Classification Test* : The battery is newly constructed and consists of thirteen tests. The validity study of this battery was conducted on 1000 students and it was found that there were a few tests in this battery which could predict college grades. High validity coefficients are observed with respect to electrical work, structural and mechanical work. A revised version has recently came out with sixteen aptitude tests.

(8) *Differential Test Battery* : There are twelve tests in this battery and these could be classified into broad seven categories. (i) Compound Series Test, (ii) General Ability Test (Verbal), (iii) General Ability Test (Numerical), (iv) General Ability Test (Perceptual) (v) Shape Judging Test, (vi) Mechanical Comprehension Test and (vii) Speed Test. Validity study of this test is done in small scale and the obtained result is encouraging.

(9) *Academic promise Test* : Seven scores, viz., (A) Abstract Reasoning (B) Numerical, (C) Non-verbal total, (D) Language usage, (E) Verbal, (F) Verbal total and (G) Total,

can be obtained from three subtests, viz., Abstract Reasoning, Numerical and Verbal included in the battery. Detailed validity data for schools of different types are given in its manual. Validity data involving prediction over three or four months are encouraging. Generally individual scores may be found more revealing than their sums.

(10) *Employee Aptitude Survey* : This battery consists of ten short and mutually independent tests. The time-limit for each of the tests except the last one is five minutes and for this last one it is ten minutes. The validity data contain 270 correlation coefficients several of them are in connection with job performance where supervisors rated their subordinates into *upper* and *lower* groups on the basis of overall job performance and using this dichotomised criterion biserial and tetrachoric coefficients are calculated.

(11) *Multiple Aptitude Tests* : This battery consists of nine tests, viz., word meaning paragraph meaning, language usage, routine clerical facility, arithmetic reasoning, arithmetic computation, applied science and machine and space relation. The technical report consists of three chapters on validity-construct, concurrent and predictive. There are sixty-three tables and forty-two graphs for occupational groups describing the validity coefficients. No multiple correlations are reported.

There are other test batteries but most of them lack proper validity study though

the results obtained with them on small-scale studies are encouraging. In addition to these American Tests discussed above there are several test batteries developed in India or adapted foreign tests to suit Indian conditions. A brief description of these test batteries is given below.

(1) *General Aptitude Test Battery* : This test was originally standardised by the U.S. Department of Labour, and in India, Singh of Gujarat University is standardising it.

(2) *Vocational Counselling Test Battery (English)* : This test battery is compiled and adapted by the Bureau of Vocational And Educational Counselling of Calcutta. Eight tests have been adapted from the Differential Aptitude test battery. The full battery consists of 4 verbal, 2 numerical and 4 spatial tests. The total grades obtained by pupils in public Senior Cambridge Examination have been correlated with stanines obtained on relevant subtests of this battery.

(3) *Educational Counselling Test Battery (English)* : This battery consists of several subtests for each of the dimensions covered, e. g., there are 7 for verbal, 6 for numerical and 4 for spatial sense. This battery is being standardising by the Bureau of Vocational and Educational Counselling of Calcutta.

(4) *Predictive Battery of Differential Scholastic Aptitude* : This test battery is available in Hindi and Marathi. It consists of eight tests (i) Numerical test, (ii) Verbal. (iii) In-

ductive Reasoning (iv) Deductive Reasoning, (v) Spatial Reasoning, (vi) Perceptual Speed (vii) Finger Dexterity and (viii) Rote Memory. Correlations between test scores and school achievement were calculated and the obtained values range from +.45 to +.66.

(5) *Shreni Vinyash Pariksha* : This is an adaptation of the Differential Aptitude-Test Battery of the Psychological Corporation. It consists four tests in Bengali viz., Numerical Ability, Abstract Reasoning, Verbal Reasoning and Scientific Aptitude. The first and the last tests have been administered on more than 4000 students while the other two tests were taken by 1500 students. Though the tests have been administered on such a wide sample it is unfortunate that no validity study has been reported so far except in case of the Scientific Aptitude Test.

(6) *An Aptitude Test Battery for Differential Prediction* : Form the descriptions of different aptitude tests which are in use in India, it is clear that many of them are only in the developing stage and they need improvement. It can be seen even for the foreign tests discussed here that as far as differential prediction is concerned not much systematic work has been done. Consequently it was felt that a battery of aptitude tests for purposes of differential prediction if developed in India would be more helpful to our students instead of adapting one of the available ones.

Considering the need to develop an effi-



cient tool for making differential prediction of success at the class VIII level at the end of which the students are to be allocated to different streams of study it. Mukherjee (1965) has constructed a and standardised a test battery for that purpose in Bengali, language. This test battery consists of ten different tests, viz., (i) English Spelling, (ii) English Usage, (iii) English Comprehension, (iv) Matching, (v) Coding, (vi) Abstract Reasoning, (vii) Verbal Reasoning, (viii) Mathematics Knowledge & Aptitude, (ix) Scientific Knowledge & Aptitude and (x) Mechanical Comprehension. These tests were twice revised on the basis of item analysis until they reached the final form.

In order to find out the differential validity of the final version of this battery, it was administered on a group of 1048 students reading in class VIII in different schools at Calcutta. The academic performance of these students were recorded upto the Higher Secondary Examination which was held after 3.5 years. The students upon whom the tests were administered at class VIII were diverted mainly into three streams of study, viz., Science, Humanities, and Commerce. Separate regression equations were developed for these three streams with the Higher Secondary Examination marks as the criterion after dividing the students sex-wise. The multiple correlations ranged from +.55 to +.71. All these coefficients were significant except one which was very close to the significant value. Reg-

ression equations were also developed by taking the school examination marks as criterion and the multiple correlation as high as +.88 was obtained there. Expectancy tables showing the relation between aptitude scores and academic performance at different ranges of the aptitude scores were developed and the results were encouraging.

### Conclusion :

From the above discussion, it is clear that ample research work remains to be done about the predictive and differential validity of the numerous aptitude tests developed so far.

One point may be added here that in the usual type of validity study only the aptitude tests are taken into consideration, But while predicting academic success one should not rely upon the results obtained with aptitude tests only. Past record on achievement tests, (usually the school examination marks) should also be taken into account. It is observed in another study by the authors with the aptitude test battery for differential prediction that validity coefficients are significantly increased if school examination marks are considered along with the aptitude test scores Chatterji and Mukherji (1970). It is also perhaps true that interests and personality traits should also be taken into account to increase the accuracy of prediction. Further research should be done in this direction,

## REFERENCES

- Bingham, W.V.D. : *Aptitudes and aptitude testing*. National Occupational Conference. Harper Brothers : New York 1937
- Buros, Oscar : *Mental measurement yearbook* 5th & 6th volumes. The Grybon Press, High Land Park, New York, 1959.
- Chatterji, S. and Mukerjee, M. : Utilisation of Aptitude Test scores in connection with school examination marks in improving allocation process at the end of Class VIII—A validity study. Accepted for Publication in *Jr. of Psychological Research*, 1970.
- Helmstadter, G. C. : An empirical comparison of methods for estimating profile similarity. *Educ. Psychol. Mnt.*, 17, 1, 1957.
- Mukerjee, M. : *Construction and standardisation of a differential aptitude test battery*, D. Phil. thesis, University of Calcutta, 1965.
- Mental Measurement Hand-book for India, National Council of Educ. Res. & Trg., New Delhi.
- Super., D E. : *Appraising vocational fitness by means of psychological testing*. Harper Brothers, New York., 1962.
- Rao, C. R. : *Linear statistical inference and its application*. John Wiley & Sons. Inc., N. Y., London, Sydney, 1965.
-