CHAPTER II

APTITUDE TESTING AND THE DIFFERENTIAL APTITUDE TESTS

2.1 Introduction

In the last chapter it was described how the conceptual framework of the term "aptitude," was already implicit in the knowledge of individual difference. The development of the idea through the ages was also briefly traced. The 19th century had brought with it the scientific attitude which emphasized experimentation and measurement. These were first applied to the physical sciences but later also attracted the attention of developing behavioural sciences including psychology. This led to the attempts to describe, define and measure intelligence and stimulated a similar interest in the measurement of aptitudes or inner potentialities. It was hoped that, if these could be measured, it would go a long way in reducing the wastage of manpower and talent. It was evident that fitting of jobs to right persons and vice versa was a subject in which people from all ages were interested. They were, however, handicapped

by the lack of objective and precise tools of psychological measurement. On account of the growth of scientific methodology and applied statistical methods these tools are now available.

The measurement activity got a big fillip in World War II, when several recruits from "fighter pilots to cooks" were to be selected. This occasion was instrumental in the development of several aptitude tests including some multi-factor tests batteries, which aimed at presenting a whole picture of an individual's basic abilities. There were tests of general and specific aptitudes e.g. numerical, artistic, mechanical dexterity and comprehension, clerical etc. Aptitude testing was now considered as a comprehensive term for assessment of abilities which could be developed with necessary training. The verbal intelligence tests, were now more appropriately termed as "scholastic aptitude tests" inasmuch as what they measured was more a capacity to learn from school training than the native intelligence.

It is not possible here to refer to the specific tests in each area, which have been developed so far. However, a fair description of such important tests can be had from several well-known books on testing such as those by Anastasi,[1] Cronbach,[2] & Freeman,[3] The various Mental Measurements Yearbooks

1. A.Anastasi, Psychological Testing.
2. L.J.Cronbach, Essentials of Psychological Testing.
3. F.S.Freeman, Theory and Practice of Psychological Testing.

and Tests in Print edited by Buros[1] are valuable sources of such information. Bingham[2] was one of the first to initiate interest in the area of Aptitude testing.

There has ~~been~~ not been much work in the field of aptitude testing in India. Most of the output has been in the field of intelligence testing, particularly verbal tests. An interest in the aptitude testing has developed recently after the recommendations of the Secondary Education Commission that Multipurpose schools be established offering opportunities to individuals to choose various technical courses according to their "aptitudes." It was felt that standardized tests can help in proper selection and allocation of pupils to various courses. The interest grew with the establishment of the Central Bureau of Educational and Vocational Guidance in Delhi, and several State Bureaux.

Much of the work in these Bureaux is being done in the field of educational and vocational guidance. Some of these centres took up some research projects for preparation of new tools of measurement. Here also, most of the work was limited to adaptation or normative studies in the region. Some normative studies on the various tests of the DAT will be referred later. As far as the investigator knows there have not been many independent reported studies in this direction. While

---

1. O.K.Buros (ed.), 3rd, 4th and 5th Mental Measurement Yearbooks, and O.K.Buros (ed.), Tests in Print.
2. W.V.D.Bingham, Aptitudes and Aptitude Testing.

some of the tests (published or unpublished) have been developed in independent research, many others developed only departmentally[1] by any government or semi-government organisation. Such tests are ordinarily not sold and hence, their use is restricted to the respective departments only. Some notable attempts in the first category are (1) A Mechanical Aptitude Test by Sharma,[2] (2) A Science Selection battery by Roychowdhury,[3] (3) A predictive battery of Differential Scholastic Aptitude, by Verma,[4] and (4) A Multi-factor test battery by Mitra.[5]

2.2 Differential Testing and Multi-factor Tests

"One of the chief distinghishing feature of contemporary psychological testing is its differential approach to the measurement of ability."[6] Considerable improvement in the methods of testing and its theory came, as stated above, after World War II. This was further promoted by the need of the times and the development of new and improved psychometric techniques. Several aptitude tests were developed for specific skills. It was felt that a battery of tests giving a set of relative scores, about the same individual on different

---

1. for example, those developed by a. Directorate General of Employment and Training, Ministry of Labour and Employment, and b. some Guidance Bureaux.
2. Unpublished Ph.D. Thesis, Agra University, 1962.
3. published with an experimental manual, published by Manasayan, Delhi, consists of three sub tests: scientific comprehension, spatial ability and numerical ability.
4. published by the University of Nagpur, Nagpur, infra.
5. unpublished, infra.
6. A.Anastasi, Psychological Testing, p. 337.

abilities and thus providing a more complete picture about him, could be more useful for occupational and educational guidance than an aptitude test to measure a specific skill.

Several factors led to this trend: first of all the fact of intra-individual variation in performance in different tests.[1] A test of general intelligence consisting of different types of items could not serve this purpose, as it brought out only a single score for interpretation. A single score, however, meaningful it might be, could not tell much about an individual's ability on different types of performances. Moreover, various subjects were to be consistent with the test as a whole and mostly they were selected so that they correlated well with each other. Thus the sub-tests in such cases were meant to "minimize, rather than maximize intra-individual variations."[2]

A further need of such Multi-factor tests was felt on account of the developing interest in educational and vocational guidance, for which purpose also an integrated battery giving a profile of an individual's different abilities was necessary. It was also thought desirable for proper comparison of an individual's performance on various tests, that such Differential tests be standardised on an identical population. Assembled batteries, developed in different contexts, may not be as effective instruments as those developed under at least a comparable

---

1. ibid., and also referred in Freeman, op. cit., p. 412.
2. A. Anastasi, loc. cit.

situation. For the "tests...standardized on different populations at different times and for different purposes, one is making a considerable interpretive leap if he assumes that the rating of 'average' has the same meaning for that student in all... areas of ability."[1]

"Differential" according to the definition given by English and English,[2] is "a distinguishing feature which makes something different from other things," and a battery of Differential Tests therefore, is a battery of tests designed to serve this purpose of presenting a profile of one's abilities, relative to his own and to a specific group. "It is noteworthy that the modern trend is away from single score intelligence tests in favour of tests which yield several scores - e.g. verbal, numerical, abstract, mechanical etc. This is a recognition of the fact that for specific courses or jobs, the best prediction can be obtained from more specific measurement than is yielded by an undifferentiated single 'intelligence' score or IQ."[3]

Another important factor which stimulated the growth of Differential testing was the development of Multiple-factor theories, primarily by Thurstone, followed by other

---

1. G.K.Bennett et al. Manual for the Differential Aptitude Tests, p. 3.
2. H.B.English and A.English, A Comprehensive Dictionary of Psychological and Psycho-analytical terms, p. 152.
3. A.G.Wesman, "What is an Aptitude," Test Service Bulletin, The Psychological Corporation, New York, no. 36 (1948).

psychometricians. The theory exploded the myth of general intelligence as a single entity and provided substantial statistical evidences in favour of the view that there are several different abilities which an individual may possess in varying degrees. Assessment of these different abilities in an individual, therefore, was an important need of the time for minimising the wastage of manpower in industrial selection and of talent in educational and vocational placement.

Thus, side by side with the specific aptitude tests, several such 'differential ability' tests'[1] were also rapidly developed. In 1957, Super edited a special series for the American Personnel and Guidance Association where important tests batteries were fully described by one of the authors, and he offered his comments with a view to assess the current status of these batteries.[2] Super expressed his admiration for these attempts towards the preparation of such tests and saw in them a new effort to meet the practical needs. He says, "this (publication of several multi-factor test batteries) represents something new, for during the generation preceding this decade factor analysis had served solely as a tool for researchers interested in the structure of mental abilities and human traits. Its practical use has only been recent."[3]

---

1. These are also known as Multi-factor tests or Multi Aptitude tests.
2. D.E.Super (ed.), Use of Multi-factor Tests in Guidance (and also Pers. & Guid. Jour., 1957)
3. ibid., p. 1.

Super mentioned four desirable qualities of a good guidance test which were his guidelines for reviewing the existing tests. These four qualities were:

1. that tests should describe; that is, they should be descriptive of the individual tested and provide sufficient information about his abilities,

2. that tests be timeless; that is, they should stand the test of time and are useful for a long period,

3. that tests should predict; and

4. that they should be multi-potential; that is, they may be applied in a general way to great many occupations taking into consideration the polygonal nature of the man as well as occupation.[1]

He included the following 8 tests used in U.S.A. at the time (and probably there has not been any notable addition up to this time):

i. Differential Aptitude Tests (DAT)
ii. General Aptitude Test Battery (GATB)
iii. Holzinger-Crowder Unifactor Tests (HCUT)
iv. Factored Aptitude Series (FAS)
v. Multiple Aptitude Tests (MAT)
vi. Flanagan Aptitude Classification Tests (FACT)
vii. Tests of Primary Mental Abilities, and (PMA)
viii. Guilford-Zimmermann Aptitude Survey. (GSAS)

The description of each test was given by one of the authors and was followed by the comments by Super. It is not possible here to fully describe them, but it may suffice to

---

1. ibid., pp. 3-4.

mention that

1. most of the batteries have striven to achieve factorial purity in the sense that the tests in the battery are measures of one major ability. Such batteries are FACT, GZAS, HCUT and FAS while others have not so much banked upon the factorial purity. One of such tests is the DAT, whose authors were mainly concerned about the production of "an integrated battery of aptitude tests designed for educational and vocational guidance in junior and senior high schools."[1] They did not strive for factorially pure tests, and in fact included some, such as Mechanical Reasoning Test, which were factorially complex but capable of high prediction for the purposes for which these were constructed viz, for the educational and vocational guidance in schools.

2. Some of the batteries were expressly constructed to serve the war time conditions. In this category may be included FACT and GZAS which have gathered the war time experiences to extend the batteries for normal use. This fact must be kept in view for a proper understanding of the nature of sampling and the consequent study which was originally based on military personnel.

Of these eight batteries, Super recommends the two as meeting the necessary standards set by him, the DAT and the GATB, the former prepared specifically for educational and

---

1.The Psychological Corporation, Catalogue 1964, p. 8.

vocational guidance, and the later for occupational guidance and selection. His comments for the DAT are: "The DAT may be characterized as currently the best battery for use in educational guidance in high school..."[1] Coughlin compared the DAT with the PMA of Thurstone and concluded that "the DAT emerges as a superior instrument for purposes of educational and vocational guidance."[2]

Aptitude Testing in India.-- In India, the aptitude testing, began only recently. Some important tests have been mentioned earlier of which two Differential batteries will be briefly described here:

1. 'A Predictive Battery of Differential Scholastic Aptitude' by Varma: Primarily prepared for educational guidance, this does not claim to have been based on a study of factor analysis though factorial loadings were reported and 4 factors were extracted. The Battery consists of nine tests: (1)Numerical, (2) Verbal, (3) Inductive Reasoning, (4) Deductive Reasoning, (5) Spatial, (6) Perceptual Speed, (7) Finger Dexterity, (8) Rote Memory, and (9) Physical relations. Total Testing time is 161.5 minutes. Reliability ranges from .60 to .93, but the manual does not report the method used. Separate validity coefficients of each test with each subject is not

---

1. Super, op. cit. p. 88
2. G.J.Coughlin, A Preliminary Investigation of the Differential Aptitude Tests. Unpublished master's thesis. Fordham University, 1950.

reported but 'corrected' validity coefficients between a group of tests and each of the 4 subjects (Language-Marathi, Geography, Science and Mathematics) was .56, .47, .45 and .66 respectively.[1]

2. A Multifactor Test Battery by Mitra: Consists of 11 tests, (1) General Knowledge, (2) Matching, (3) Classification, (4) Following directions, (5) Perception, (6) Dexterity, (7) English Usage I, (8) English Usage II, (9) Simple arithmetic, (10) Greater and lesser numbers and (11) Number series. Total testing time - 72 minutes. No reliability was reported but a complete factor analysis using Thurstone's Centroid method was done extracting two factors. Factorial validity was reported. It was prepared primarily for selection of entrants in statistical and other courses in Indian Statistical Institute.[2]

The former test is published but is not available for sale. The later is unpublished so far.

2.3 Differential Aptitude Tests

Differential Aptitude Tests, or the DAT as they are popularly known, is a multifactor test battery of eight aptitude tests.

---

1. Ministry of Education, Govt. of India. Report on a Predictive Battery of Differential Scholastic Aptitude, University Training College, Nagpur, 1958.
M.Varma, Manual for the University Training College Battery of Tests.

2. S.K.Mitra, A Multi-factor Test Battery, Education and Psychology, Vol. 5, no. 3 (September 1958) pp. 147-157.

Verbal Reasoning	(VR)
Abstract Relations	(AR)
Space Relations	(SR)
Numerical Ability	(NA)
Clerical Speed and Accuracy	(CSA)
Mechanical Reasoning	(MR)
Language Usage - Spelling	(LU-sp) and
Language Usage - Grammar	(LU-gr).

DAT and the letters in parentheses signify the abbreviated names of tests; by which the tests would be referred in the entire report which follows.

Unlike several other multifactor test batteries, described in the preceeding section, these are not uni-factor tests that is, tests are not based upon any systematic factor analysis each test a measure of a pure ability. The rationale is very simple: the authors were motivated by the desire to develop a battery of tests which could differentiate between individuals, and could predict to some extent the academic or vocational success of the individual. As Super comments, "the emphasis is on utility as contrasted with ~~comment~~ factorial purity."[1] ". . . in some instances (authors) deliberately sacrifice factorial purity for predictive validity."[2] This seems to be the practical purpose, unburdened by any theoretical consideration though the various factorial studies, especially those of Thurstone, seem to have influenced.[3] In the words of the authors, "the Differential Aptitude Tests were developed to provide an integrated, well standardized procedure for measuring the

1. Super, op. cit., p.20.
2. ibid., p.88.
3. Super, Apprasing Vocational Fitness., p. 130.

abilities of boys and girls in junior and senior high school. Since it was clear that no practical test battery could encompass measures of all aptitudes, the attempt was made to include tests which would be useful in many areas and would yield scores which were directly interpretable by informed counsellors. Inherent in the conception was the conviction the tests must prove themselves in actual use to be effective predictors of how well students perform.[1]

The Differential Aptitude Tests were prepared by Drs. G.K.Bennett, H.G.Seashore, and A.G.Wesman and were published by the Psychological Corporation, first in 1947. The battery of tests has been "carefully constructed and impressively standardized."[2] In itself, the project was a stupendous one. The original manual of 1947 reports standardization of 20,000 boys and girls in 30 school systems of U.S.A., and in 1952 revised norms were published which were based on 47,000 students in over 100 schools systems in 26 states of the country.[3] "The amount of validity data available on the DAT is overwhelming, including several thousand validity coefficients."[4] In addition, two long range follow-up studies were reported.[5] The details of various validity and normative studies have been accumulated together and is available

---

1. G.K.Bennett et al, The Differential Aptitude Tests: an overview, (quoted in Super, op. cit., p. 9).
2. G.J.Coughlin, op. cit.,
3. G.K.Bennett et al, loc. cit.
4. A.Anastasi, loc. cit.
5. G.K.Bennett et al, Manual for the Differential Aptitude Tests, pp.

with the American Documentation Institute in form of micro-films or micro-cards.[1]

It is not definitely reported anywhere how the tests were finally selected nor the present investigator could have any access to original documents about this. It may, however, be presumed, after going through several source materials, that these eight tests were chosen on basis of experiences gathered upto that time including factorial studies. Bennett,[2] the senior author, refers to the genesis of these tests in one of his papers. Because of some interesting light this throws on how the DAT were developed, the relevant expert is quoted in full:

> "...This series of eight tests was, from its inception, designed to produce a profile of scores useful in the guidance of high school students. It was apparent that the battery must have good differential power to justify its use in the manner intended and the Kelley method proved of help throughout. The first step in this development was to study available test data to see which pairs of meaningful measures could effectively be used at the same time. Since the practical upper limit of reliability for short power tests is about .9, it was decided that the maximum permissible intercorrelation would be about .6 with lower values desirable. This approach clearly indicated that certain combinations of variables, such as arithmetic problems or synonyms and verbal analogies could not be expected to contribute importantly to the diagnosis of ability. On the other hand, we were somethat surprised to find that spelling and grammar, which we had first intended to consider as subsections of "language usage," were sufficiently independent to warrant individual consideration. Having selected a promising set of tentative variables, the next step was to find item forms suitable

---

1. G.K.Bennett et al, Supplementary Tables for the Differential Aptitude Tests. Washington: American Documentation Institute, 1952, no. 3656.

2. G.K.Bennett, Evaluation of Pairs of Tests for Guidance Use. Paper read before the Amer. Psychol. Assn., 1947.

to the ability levels intended, namely grades 8 through 12, and producing the maximum reliability per unit of time. It is not feasible to discuss here the various item types tried, but we have succeeded in producing a series of 8 tests, none over 30 minutes in length with reliabilities ranging between .85 and .93 for a single grade of high school boys and, with the exception of the Mechanical Reasoning Test, between .86 and .92 for a single grade of high school girls. The intercorrelations fall between .06 and .62 for boys, and between .12 and .67 for girls. The respective mean intercorrelations are .38 abd .39."

The eight tests are now described:

1. Verbal Reasoning.-- This is the test of verbal comprehension and is a "measure of ability to understand concepts framed in words."[1] This test employs analogy type items in which two words, the first and the last, are to be filled in from the 5 pairs given as the alternate responses (in Form L). The items are universal and have been drawn from different areas of knowledge, e.g. geography, history, politics, science etc.

The analogy type of items is particularly useful as a measure of reasoning ability. In most of verbal intelligence tests this is an important item; in fact, the situation presented by such situation brings into play complex intellectual (reasoning) processes. This type of items stimulates one to seek and to know further and to "educe relations," which is one of three neogenetic laws of Spearman according to which the mind "creates new mental content."[2] The analogy

1. G.K.Bennett et al, Interpretive Manual, [illegible] p. 6.
2. J.C.Flugel, A Hundred Years of Psychology, p.264.

type items have been most fundamental in intelligence testing, and have been used from Binet onwards to the modern tests of intelligence in one way or other.

The VR test is thus designed to predict success in areas where complex verbal relationships and concepts are important. Besides success in academic courses, this may also indicate "something of the occupation level" since there is a positive relationship in many occupations between the level of responsibility of a job and the complexity of verbally phrased items to be comprehended.

In the Form L, this test consists of 50 such items each in the form of a sentence of which the first and the last words are blank but are to be found from among the five alternate pairs given. Form A, also presented the same types of items but there were 16 possible combinations. The time limit for the original Form L test is 30 minutes.

2. Numerical Ability.-- Like the VR, this test is also a measure of the ability that is indicative of general learning ability; specifically, it is an essential ability in such courses as Mathematics, Statistics, Book-keeping and such occupations as clerical, carpentary, engineering etc. Computation type items rather than reasoning type items have been employed to bring in the minimum verbal element.

The authors have tried to frame the items in a way which stimulates independent thinking. Thus, in one item, two

answers are 47ft. 24ins., or 49ft. both of which are correct but the correct score is earned only if the later is marked right. Necessary instruction to think this way are given. There are in all 40 items with time limit 30 minutes.

3. Abstract Reasoning.-- This is a non-verbal measure of the reasoning ability and like Verbal Reasoning, is also a fundamental ability of the general intelligence. The general reasoning process is not much different from that of the VR but in this case the pupil is not handicapped by the lack of proper knowledge of the English language. The student has to perceive relations and has to think in abstract symbols. Difficulty level as also the conceptual complexity increases gradually. "The differences are apparent; discerning why the patterns differ is the intellectual exercise."[1]

The test has fifty items, each consisting of 4 problem figures which depict a relation between the first four shapes and the student is asked to locate the correct shape from among the five given as 'answer figures' which will fit in as the fifth shape of the problem figures. The ability is an important one for general intelligence as well as also for vocations or school courses which demand such ~~prepa~~ perception of relationships in symbols. The time limit is 25 minutes.

4. Space Relations.-- This is an ability for the [illegible]

1. Bennett et al. Manual p. 7.

occupations where structural visualization in an important requirement such as, architecture, mechanical drawing, civil engineering, dressmaking, decoration etc. Spatial ability, though is not correlated well with usual school courses, is recognized as an important one for prediction of success in some technical courses, including mathematics. Smith quotes Reese Edward as stating that "where selection is for technical education alone...spatial tests have been shown to have considerable value."[1] Smith,[2] in his excellent book, has gathered several investigations and quotes the report of Yates and Pidgeon who say that though "such tests are not usually considered as useful predictors of success in grammar school, it is possible that the abilities measured by this kind of test are related to subsequent success in some brances of mathematics and science."[3]

This test employs two approaches to spatial visualization: ability to visualize from a given pattern and ability to imagine how it would appear. "A feature inherent in these items is that they require mental manipulation of objects in three-dimensional space. Item forms which refer to only two dimensions are less useful, since there are relatively few occasions when perception of two-dimensional space alone is important."[4]

---

1. R.Edward, quoted in Smith, Spatial Ability, p. 33.
2. I.M.Smith, Spatial Ability.
3. Yates and Pidgeon, quoted in Smith, op. cit. p. 29.
4. Bennett et al, Manual, p. 7.

There are 60 such problems (in form L; there were only 40 in Form A) and the drawings are large and clear so that visual clues and slight discriminations do not affect the total score of the individual. The time limit is 25 minutes.

5. Mechanical Reasoning.-- The test consists of pictures depicting some situation or activity, where some mechanical and physical laws come into play. The student is asked to understand this and has to mark one of the few alternate responses. It is clear that the test demands understanding of simple physical laws. This is a measure of the ability to perceive simple mechanical relationships, which may be regarded as an aspect of intelligence, if intelligence is broadly defined.[1] The performance is affected by formal training and previous experience "but not to a degree that introduces serious difficulties in interpretation."[2]

The test is a revised series after the model of Mechanical Comprehension Test by Bennett, and is equal in difficulty level to the Form AA of the TMC. As this is a later development, the authors claim considerable confidence in this test. "The test is useful in those curricula and occupations where an appreciation of the principles of common physical forces is required...Occupations such as those of carpenter, mechanic, maintenance man, assembler, and hundreds of others in plants

1. ibid., pp. 7-8.
2. ibid.,

and factories require the kind of understanding it measures."[1] It is a test definitely for vocational guidance and its use in educational guidance is limited as is apparent from the low validity correlations with usual academic courses. The test consists of 68 such items and the time limit is 30 minutes.

6. Clerical Speed and Accuracy.-- This is a pure speed test, which is an important factor in routine clerical work. It measures "speed of perception, momentary retention and speed of response"[2] and involves no or only little intellectual complexity. It was found from observation that "errors are rarely made in a task as simple as this one."[3]

The ability is not so important educationally but is important for such occupations as checking, filing, coding etc. where the speed of a performance is an essential requirement. It is possible, however, that as the speed of perception, neatness and precision are important accessory qualities of good student, the score on this test may reflect the students' care for such work.

"Clerical tests are as applicable to the guidance ~~of a~~ problems of a school as they are to the personnel problems of a business organization."[4] They are useful in spotting individuals who perform poorly in clerical tests for giving

---

1. ibid.,
2. Bennett et al. Manual, p. 8.
3. ibid.,
4. G.K. Bennett and Ruth M. Cruickshank, A Summary of Clerical Tests, p. 68.

them training in line with their abilities.[1] Good students generally, perform well on this test, though in the first instance their performance may not be equal to their performance in other tests. This is, probably, due to their stress on doing correct rather than on the speed. In an experiment by the authors, some students who scored low at the first testing, scored quite high when tested another time after the need for fast marking was re-emphasized.[2]

The test has two equivalent parts, each containing 100 items. Each item consists of some combinations of letters or figures is given of which one is underlined. The student is required to locate this underlined combination in the answer-sheet in which same combinations as in the items are given but in a different order. The time limit is 3 minutes for each part of 100 such items.

7. Language Usage-spelling,-- and
8. Language Usage-grammar,-- (LU-sentences in Form A).--

Out of the entire series, these two are perhaps nearest to achievement tests. These are the measures of two distinct and independant abilities, usually thought to be one. In a research quoted by the author[3] it was found that high or low score on one did not necessarily mean an identical score on the other. These two are, however, abilities which often go

---

1. ibid., p. 70.
2. Bennett et al. Manual, p. 8.
3. G.K.Bennett, Manual, p. 9.

together and a consideration of both scores may be a good indicator of educational achievement (in English language) and even might be a good diagnostic instrument indicating the need for a remedial action. For the sake of economy and convenience, these two tests have been combined into the same booklet.

The spelling portion of the test consists of 100 words, both correctly and incorrectly spelled, which have to be so recognized by the student who marks 'True' or 'False' for each word. The wisdom of using correctly spelt words may be questioned but experimental study by some of the authors showed that properly chosen correctly spelt words may be reliable and discriminating items.[1] The authors also quote a study by Cronbach which "resulted in the discovery of enough words for the test which were effective items when correctly spelled."[2] The list of words were obtained from the Gate list of frequently mis-spelled words.[3]

The grammar portion of the test (Form L)[4] has 60 sentences, each divided in four parts. Both correct and incorrect sentences have been used. In case of a grammatically wrong sentence, the error would be only in one of the four parts (indicated by the letters A, B, C, D). The student has to

---

1. A.G.Wesman, The Usefulness of correctly spelled words in spelling test, J. Educ. Psychol. pp.242-246.
2. quoted in Bennett at al. Manual, p.8.
3. ibid.,
4. There were only 50 items in Form A of this test.

recognize the error and has to mark the corresponding letter on the seperate answer sheet. In case of correct sentences, the student has to mark the fifth alternative, the letter E. The time limits are 10 minutes for the spelling and 25 minutes for the grammar test.

Characteristics of the DAT battery.-- Two parallel forms- Form A and B- were published originally. These forms were used for a long time and the latest Manual, 3rd rev., is also based on these. However, recently, new and revised Forms, L and M were published, though these are in essential respects identical to form A and B. The Manual for the new forms has not been published and the users are asked to use the same manual as for the A and B for these too until new Manual is ready with technical information based on the new forms. Only a booklet entitled "Directions for Administration and Interpretation" for forms L and M is now available which gives the revised norms.[1]

The form L (and M) differs from form A (and B) only in certain respects, one of which is that eight tests are now published in two booklets, each containing four tests, with one answersheet for all the four. This brings about an economy in the use of tests, especially when they are to be used in large number, and to some extent a simplifies the testing procedure. In addition to this, the format of some of the

1. Published by the Psychological Corporation, 1963.

tests has been changed. There is no change in AR, MR, NA and CSA which remain same as Form A; in others the major changes are:

1. slight change of the number of items in SR and LU-gr. The revised number is 60 items in each while there were only 40 and 50 respectively in SR and LU-gr, Form A.
2. making only one correct response in the above two tests, i.e. SR and LU-gr. (In Form A, the correct responses were more than one in some items), and
3. giving five definite pairs of possible answers in case of the VR, from which one pair is correct. (In Form A there were 16 possible combinations)

The time limit of the various tests also remain as in Form A, except that in SR, where the revised time limit is 25 minutes (instead of 30 in Form A). The total testing time for all 8 tests is now 181 minutes. (The total testing time for the present revision is 196 minutes - 15 minutes more than the original testing time in the U.S.A. It shall be discussed later in this report)

Much is not yet known, upto the time of writing this report, about the rationale of change in this revised Form L of the DAT battery. The relevant technical information of Form L (and M) is also not yet published. As for practical purposes, the technical data of the Form L remains the same

as Form A. The following is the relevant information about the Form A:

Reliability.-- The reliability of all the tests, except CSA which is a speed test, was studied by split-Half technique. Reliability for the CSA was studied by the Equivalent Forms method, being the correlation between the scores of Forms A and B. Reliability was studied for each grade from 8 to 12 (of American schools), seperately for each sex. The average Reliability coefficients (for all grades) ranged from (1) .85 to .93, in case of boys and (2) .71 to .92, in case of girls. These are shown in Table 1 below:

TABLE 1

Average Reliability Coefficients for the Differential Aptitude Tests

	VR	NA	AR	SR	MR	CSA	LUsp	LUsent*
Boys	.90	.90	.90	.93	.85	.87	.92	.88
Girls	.90	.86	.89	.90	.71	.87	.92	.87

* In Form A, LUgr. was known as LU-sentences.
SOURCE: From Bennett et al, DAT:an overview, in Super, Use of Multi-factor tests in guidance, pp.9-19.

The reliability is also presented in terms of the Standard Errors of Measurement, the median values of which range from 2.8 to 6.4, in case of boys and 2.6 to 6.6, in case of girls.

Intercorrelation.[1]-- The intercorrelation of various tests

1. Bennett et al, Manual, pp. 67-70. Also in Bennett et al, DAT: an overview, in Super, Use of Multi-factor tests in guidance, pp. 9-19.

of Form A, ranged from 0.6 to 0.67, and the median inter-correlation coefficient was 0.41. 'The proportion of differences in excess of chance proportion' was also studied. The criterion value of 25% was exceeded by all the pairs, in case of boys, and by 24 pairs (out of 28 possible pairs) in case of girls.

Validity.-- The validity data are overwhelming; about 4000 studies have been reported which are available in form of micro-cards or microcopies from the American Documentation Institute. It is not possible here to quote the original studies, but few observations may be made:

1. Predictive validity was studied for school success. The criteria were (i) the school grades in various subjects, and (ii) standardized achievement tests.

2. Concurrent validity was also studied with scores on achievement, aptitude, and intelligence tests and Kuder Preference Record.

3. Two long follow-up studies after the first testing in high school, have been reported. Such studies serve the two purposes: firstly, they indicate the stability of the tests themselves by indicating the correlation between earlier and later administration and secondly, the studies reveal important traits and shifts in occupational analysis. Of the two studies reported, one[1] is a longitudinal study 4 to 5 years after the

1. E.Doppelt and G.K.Bennett, "A Longitudinal Study of the Differential Aptitude Tests," Educational and Psychological Measurement, 11:220-237 (Summer 1951.)

first testing in 1947 and another is a 7 year follow up in 1954-55.[1]

Besides the above results, several studies by the authors themselves and a number of validational studies by several other researchers have been reported in various publications. Seashore has made a comparative study of the DAT with the other well-known tests viz, CEEB Scholastic Aptitude Test, and ACE Psychological Examination which are used mainly as College Entrance Examination. He has concluded that scores on the DAT are valuable for predicting the performances for College Entrance Examinations from the CEEB-SAT.[2] Layton[3] refers to a study made by him for finding the relationship of IXth grade DAT scores to that of XIth class.[4] Harris refers to his study of DAT scores as related to grades in class XI on a seperate group. He concluded that, "granting limitations, authors contend they have established that DAT battery probably can be very helpful in predicting early in high school acceptance or rejection by the University of Hawaii and approximate college academic performance. We can help school counsellors to

---

1. G.K.Bennett, "The DAT-A Seven-Year Follow up," Test Service Bulletin, no. 49, The Psychological Corporation. Also in Bennett et al, DAT: an overview, and Bennett et al, Manual, p. 77.

2. H.G.Seashore, "Tenth Grade Tests as Predictors of Twelfth Grade Scholarship and College Entrance Status," Jour. of Counseling Psychology, I:106-115 (1954).

3. W.L.Layton and E.O.Swanson, "Relationship of ninth grade Differential Aptitude Test scores to eleventh grade test scores and high school rank," Jour. of Educational Psychology, 49:153-155.

4. Y.Y.Harris and A.A.Dole, "A Pilot Study in Local Research with the Differential Aptitude Test Battery," Personnel and Guidance Jour., 39:127-132 (Oct. 1960).

identify the worst and best bets for higher education."[1] Layton[2] also concluded that the "DAT was perhaps most useful in predicting school success." A study by Doppelt reports the validation of the DAT for auto mechanics.[3] Stoughton[4] studied the predittive validity of DAT scores for technical schools. A study by Coughlin,[5] already reported earlier, established DAT as a superior instrument for use in high school guidance.

Above are only few of the many studies done on the DAT and their predictive value. All this indicate that this battery is "perhaps the best battery for use in educational guidance in high school."[6] It was a"major psychometric event"[7] and is highly regarded as one of the best available differential batteries in use today.

DAT in India.-- The DAT have been widely used in India, particularly in the past few years. Though its use was limited only to the five non-verbal tests (the three entirely verbal tests were not widely used), it was being recognised as one of the best available test batteries for India. The use, however, was handicapped by the import difficulties and the high

---

1. ibid.,
2. Layton, Loc. cit.,
3. E.Doppelt et al, "Validation of the Differential Aptitude Tests for auto mechanics and Machine Shop students," Personnel and Guidance Jour., pp. 648-655 (May 1959).
4. R.W.Stoughton, "The differential predictive values of the Differential Aptitude Tests in the Connecticut technical schools" unpublished Doctor's dissertation, Univ. of Connecticut, 1955.
5. Coughlin, op. cit.,
6. Super, Use of Multi-factor Tests p. 88.
7. by an anonymous psychologist, quoted in Coughlin, op. cit.,

cost. This led to the printing of an Indian edition of the entire battery by a special agreement between the original publisher, the Psychological Corporation and an Indian publisher, Manasayan. This was done with the view to make the tests easily available at a lesser cost. The entire series of 8 tests of Form A was published in 1961. In the Indian edition, some minor changes were made without effecting the actual test content, to make the tests more suitable for Indian pupils. These were incorporated after the experiences of the tests users in India, and some studies.

The major changes were (1) slight re-wording of the directions in some of the cases, (2) putting two additional examples in the AR, (3) re-wording of some questions in the MR and (4) altering the spellings in LU-sp., LU-gr. and VR test in conformity with the British spelling which was used in India, instead of the American spelling. This was done in cases where the two spellings differed, e.g. instead of American 'behavior' British 'behaviour' was used. Such cases, however, were not many. In addition to these some sentences in the LU-gr. test had to be recast from this point of view. The change in the Language tests (including the VR) were made mainly on basis of the observations of the experiences in Calcutta Anglo-Indian high schools, for these are among the schools where the verbal tests are chiefly used.[1] These

1. The present writer acknowledges the help of Dr.Riddle, who prepared the original verbal tests for India, Dr.Harper, who edited the revised series of DAT tests for India, and Manasayan, who placed the relevant records at his disposal.

revised tests were also used in Christian schools in Punjab and U.P. These three tests, due to their exclusively verbal nature were used ordinarily in schools where the chief mother tongue of most of the students was English, and also in some schools where medium of teaching was English, or where a majority of students used English as a major second language by force of habit or tradition.

The present investigation is based on the Form A of those tests which are published in India, but which do not differ in content from the tests of Form L. It was mentioned earlier that in 4 tests there was no change in Form L. Three tests (LU-sp, Lu-gr, and VR) were newly constructed in Hindi and the format followed in the present investigation is of the Form L. The items in SR are also identical to Form L.

The writer was able to gather some information about the normative studies that have been made on these tests. Ordinarily these norms are for departmental use and are not available at large. Some normative studies have been made for one or more of the DAT tests in States of Bombay, Bengal, Bihar and U.P. Study by Riddle for Punjab schools has already been referred earlier. The present investigator was able to get rather elaborate information of the normative studies in Bombay, Institute of Vocational Guidance, and Calcutta, Bureau of Educational and Vocational Counselling.[1] The reports received

1. The writer is thankful to Miss M.Ruthquist, the then Director of the Bureau of Educational And Vocational Counselling, Calcutta, and Mr.A.W.D.Costa, the then Secretary of the said Bureau who furnished the necessary details and copies of the various studies.

from Calcutta Bureau included norms for each sex, class, age and region and those from Bombay included norms for the AR, and some prediction studies for the courses of commerce, science, engineering, medicine, commercial art and ~~artie~~ architecture. Incidentally in the Bombay study[1] the organizers found 35 minutes as the most suitable time limit for the AR test. The reliability of the AR test was found to be .93. Both of these figures are identical to the results of the present investigation reported later in the appropriate place.

Some recent studies include (1) normative studies for the Ar test, by the Central Bureau of Educational and Vocational Guidance, Delhi, and (2) Normative studies in the schools of U.P. by Manovigyan Shala U.P. (Bureau of Psychology, U.P.), Allahabad for several DAT tests.

2.4 Summary

The interest in measurement of intelligence which was originated by the overall scientific advancement was further extended to the assessment of the aptitude through psychological tests. The practical need of the times i.e. placement of an individual for right type of training and thus to reduce the wastage of human resources, provided an incentive for aptitude testing. A big occasion for large scale testing, and experimentation was provided by the World War II.

---

1. Institute of Vocational Guidance, Bombay, The DAT: Abstract Reasoning Test (mimeographed).

Work in this direction, coupled with the studies in factor analysis resulted in the construction of Multi-factor Test batteries which provided a profile of an individual's various aptitudes. Some important batteries were mentioned with special emphasis on the DAT, the starting point and the main subject matter of the present investigation.