CHAPTER VI

TEST CONSTRUCTION AND ADMINISTRATION OF THE TRYOUT TEST

CRITICAL STUDY OF THE LITERATURE ON TEST CONSTRUCTION

The investigator got some experience of test construct-ion when he constructed a 'Diagnostic Test in Basic Arithmetic Skills' for his M.Ed. Dissertation. This work initiated him in test construction. Encouraged by this successful attempt, he selected the present topic for his Doctorate Thesis. The topic, naturally, calls for more detailed and scientific information regarding mental tests and mental-test-construction.

TYPES OF MENTAL TESTS

Following is a classified list[1] of most of the mental tests available in the market. The main headings are:

I. Attainment, Achievement, Diagnostic or Educational Tests.

A. Examinations.

---

1 Vernon, P.E. "The Measurement of Abilities", University of London Press Ltd., London, 1956, pp. 153-154.

B. Standardized Educational Tests.

1. Individual reading tests.
2. Language development and Vocabulary.
3. Individual oral arithmetic.
4. Group tests.

II. Individual Intelligence Tests

A. Versions of the Binet-Simon scale.
B. Wechsler scales.
C. Miscellaneous scales.
D. Performance tests

III. Group Intelligence Tests.

A. Oral
B. Non-verbal
C. Verbal.

IV. Aptitude and Special Ability Tests.

V. Personality Tests and Assessments.

Any person desiring to get thorough information of the literature on Mental measurement should refer to the Mental Measurements Yearbooks, edited by Oscar Krisen Buros. The investigator, in the beginning, went through the following volumes of 'Mental Measurements Yearbook'.

1. The Nineteen thirty-eight Mental Measurements Yearbook,

2. The Nineteen forty Mental Measurements Yearbook,

3. The Third Mental Measurements Yearbook,

4. The Fourth Mental Measurements Yearbook, and

5. The Fifth Mental Measurements Yearbook.

Besides these volumes, he went through the revised edition (1956) of 'Encyclopedia of Educational Research ' - a project of The American Educational Research Association edited by Walter S. Monroe, a distinguished Professor of Education, University of Illinois.

The reading of these volumes supplied the following types of information:

1. A list of published tests in different fields.

2. Comments and criticisms on the tests by experts from Educational and Psychological fields.

3. A list of books on different types of Mental tests construction.

4. A list of contributing test Reviewers, and

5. A list of Magazines, Journals, Reviews etc. discussing Mental Tests and Educational and Psychological Researches.

Enriched with this information the second step that

the investigator took was to study important books available on Achievement Tests, Intelligence Tests, Diagnostic Tests, Personality Tests etc.

The chief recently published sources of information on how to construct printed tests include books written or edited by Adkins, Guilford and Laccy, Lindquist, Stuit, Travers, Greene, Jorgensen and Gerberich, Gulliksen, Ross and Stanley, Micheels and Karnes, Cronbach, Barr, Davis and Johnson, Mursell, Anastasi Anne, Bean etc. Most of these publications include discussions of how to write and edit items and the procedures of standardizing mental tests. The study of 'Personality Tests and Assessments' and 'The Measurement of Abilities' by Philip E. Vernon and 'Mental and Scholastic Tests' by Cyril Burt also served the purpose of the test-constructor to a great extent.

The following books on Test construction by Indian authors also were studied.:

1. Bhatia, C. M., Performance Tests of Intelligence under Indian Conditions.

2. Kamat, V. V., 'Measuring Intelligence of Indian Children'.

3. Sohan Lall, 'Mental Measurement'.

4. Desai, K.G., The Construction and Standardization of A Battery of Group Tests of Intelligence in Gujarati,

5. Shukla, N. N., Individual Tests of Intelligence for Gujarati Children.

This study enabled the investigator to get a thorough grasp of the general procedure to be followed in mental test construction.

Obviously the procedures to be followed in constructing different types of mental tests cannot be exactly the same. For example, if we are to construct achievement tests in different school subjects, we are required to study the syllabi in the subjects critically, write the objectives and decide the weight-age to be given to each independent topic in the syllabus. We ultimately prepare a blue-print of the same and then start writing the items for the test. Thus we shall have different objectives for different subjects and different weightage shall have to be assigned to different topics of the syllabus for different subjects. Thus we shall have different types of blue-prints for constructing achievement tests in different school subjects. On the contrary, most of the "intelligence" tests contain very similar types of items. Apart from the basic differences associated with verbal and non-verbal and performance and non-performance tests, the contents of intelligence tests commonly include several of the following types of test items. Vocabulary, general information, verbal analogies, figure analogies, same - opposites, completion, arithmetic reasoning, arithmetic computation, number series, identification of missing parts, spatial relationship, proverbs, verbal and pictorial

absurdities, mazes, story and picture sequences, mixed words or sentences and reading comprehension.

Similarly procedures for constructing specific aptitude tests must be different from those of constructing intelligence tests because both specific aptitudes and intelligence are not similar. It is, therefore, necessary to study special books on aptitude test construction. As far as the investigator could gather information, very few books on Aptitude testing are available. From among these rare books, the investigator could procure and study the following books:

1. Super, D.E.,'Appraising Vocational Fitness', New York: Harper & Brothers, 1949.

The reader finds the most complete, detailed and comprehensive discussion of aptitude tests, their uses, interpretation of test results, and summary of research findings in the book. This book should be consulted by persons who wish detailed and authoritative information about specific aptitude tests.

2. Thorndike, R. L., Personnel Selection: Test and Measurement Techniques, New York: John Wiley and Sons, 1949.

Although the purpose of this book is broader in scope than the purpose of the Super's book, it contains comprehensive and authoritative information about aptitude tests and related

procedures in personnel selection.

3. Bingham, Walter Van Dyke, 'Aptitudes and Aptitude Testing', Harper and Brothers Publishers, New York and London.

This book is very useful to a person who wishes to construct an aptitude test. The author discusses the meaning of aptitude, the theory of aptitude and relation between Intelligence and Aptitude. The author also discusses the aptitudes for different professions. In the later chapters the procedures for construction and standardization of aptitude tests are discussed nicely. The book does serve the purpose of a text-book on Aptitude Test construction.

4. Hull, C.L., 'Aptitude Testing', Yonkers, N. Y. World Book Co., 1928.

The book is written in two parts. In part one, an account of the fundamental principles of aptitude testing is given; and in part two, an intelligible description of the most effective and the most economical methods of constructing batteries of aptitude tests is made. The book, though, written about thirty two years ago, even to-day serves the purpose of a text-book on Aptitude Test construction.

5. Sandiford, Cameron, Conway and Long, 'Forecasting Teaching Ability', Bulletin No.8 of the Department of Educational Research, University of Toronto,

Toronto-5, Canada, 1937.

The bulletin gives a nice account of very useful studies made by Sandiford and others, of factors related to 'Teaching Ability'.

A good deal of work is done on "The Prediction of Teaching Success" and similar topics by a number of investigators. The reports of their investigations have been published in the form of Bulletins or in different journals such as Journal of Educational Research, Educational Administration and Supervision and many others. The investigator could refer to some of them.

After the study of the books on test construction in general and those on Aptitude Test construction in particular, the next step was to study the tests on 'Aptitudes for Teaching'.

In chapters III and IV, we have discussed at length a number of Aptitude tests. Without undue repetition of the same, we shall discuss here some 'Aptitude Tests for Teachers' which the investigator could get and try to list the factors related to Aptitude for Teaching.

In the earlier chapter, we have already said that a lot of work has been done on this topic in the U.S.A., while practically no work is done in this country. Educational and other conditions in this country are quite different from those of the U.S.A. or Canada or U.K., and so the tests constructed

to suit the prospective teachers there, cannot be administered to the prospective teachers here. A new 'Aptitude Test for Teachers', therefore, has got to be constructed to suit Indian conditions.

Even then, the study of such foreign tests will give us some knowledge concerning the factors which condition success in teaching. After determining the factors which condition success in teaching, a technique of investigation, for finding out whether or not the factors are possessed by prospective teachers who wish to enter the profession, can be developed easily.

Another source of determining such factors is job analysis. This is discussed in later pages of this chapter.

A number of investigations have been carried out on the 'Prediction of Teaching Success' and allied topics. According to Sanford and Trump as many as 675 articles and monographs concerning various problems incident to the pre-service selection of teachers have appeared since Meriam's study in 1905.

The investigator could get a printed copy of 'Aptitude Test for Elementary School Teachers - In- Training alongwith a manual descriptive of the same and 'Forecasting Teaching Ability' - bulletin No.8 by Sandiford, Cameron, Conway and Long of the Department of Educational Research,

University of Toronto, Canada. He could not get the complete tests used by them but he could study the sample items for each test given in the bulletin.

Over and above a critical study of these two tests, the reviews and references about the other tests were also studied.

The above study revealed that generally the following types of sub-tests are included in test-batteries to measure aptitude for teaching.

Henry Bowers' Aptitude Test for Elementary School Teachers in Training consists of seven parts as follows:

(1) Opinions (attitudes towards various social activit-ies);

(2) Books (attitudes towards the titles of six imaginary works);

(3) Occupations (preference for certain occupations);

(4) Interests (expressed interests in various activities);

(5) An aspect of judgment (judgments of the degree of truth or falsity of several propositions);

(6) Performance (ratings of various traits such as resourcefulness, vocabulary, ability to organise material etc. in a standardized situation);

(7) High School Percentile (a measure of High School performance).

Thus the measures derived from the test include expressions of preferences and judgment in a paper-and-pencil test situation, ratings, and a measure of previous academic achievement.

Sandiford and others in their test battery on 'Forecasting Teaching Ability', includefollowing sub-tests:

(1)	General Information Test	(250 items)
(2)	Vocabulary Test	(100 items)
(3)	Completion Test	(50 items)
(4)	Paragraph Reading Test	(50 items)
(5)	General Science Test	(270 items)
(6)	Social Studies Test	(285 items)
(7)	Foreign Literature Test	(295 items)
(8)	Fine Arts Test	(240 items)
(9)	Bernreuter Personality Inventory	(125 items)
(10)	Questionnaire	(68 items)

Thus in the battery there are ten "paper tests" consisting of eight objective examinations, a questionnaire, and a personality test. The tests are designed to measure intelligence, knowledge of subject matter, wholesome personality, a desire for service, etc. possessed by the prospective teachers.

The traits in prospective teachers, which Bossing tried to measure included ability, discipline, industry, character, personality, personal appearance, health, loyalty and co-operation, and attitude towards the community.

Other factors studied by a great number of investigators include age, sex, experience, salary, skill in handwriting, skill in reading, height, weight, height-weight ratio, number of letters in signature, socio-economic status, sincerity, enthusiasm, interest, participation in extra-curricular activities in high school, scholarship in high school, scholarship in college, intelligence test score, personality rating, and score on various types of tests, such as those of professional information and those covering the subject matter the teacher is supposed to teach.

The researches reported indicate that only a few of these factors such as intelligence, scholarship, personality, and scores earned on professional information and subject matter tests are important.

Henry Bowers,[1] in his manual descriptive of the Aptitude Test for Elementary School Teachers - in- Training, says,

---

1 Bowers, Henry, "Manual Descriptive of the 'Aptitude Test for Elementary School Teachers - in- Training, J.M. Dent & Sons, Ltd., Toronto, 1948, p. V.

> The teacher is not judged solely by his success in the classroom, by his ability to produce desired changes in the pupils - the criterion of success used by Professor A. S. Barr and his associates in the University of Wisconsin - by his competency in the office, or his gifts on the playground. He must deal successfully with colleagues, parents, trustees, and supervisors, with the inspector, the school nurse, the janitor, the..... He must be an acceptable member of the community.

He says again,[1]

> Who would have the temerity to declare that this, this and this are the factors which determine success in teaching? We have been deluded by an apparently simple word 'teaching', the meaning of which each of us knows, a meaning which each of us assumes is the self-same meaning of every other person. "Teaching" is not the name of an occupation. It is the generic name of hundreds of occupations. There is not one kind of success in teaching, but hundreds.

The test constructor is also expected to possess adequate knowledge of statistics as he is to use it every now and then in test construction. The investigator being a student of Experimental Education, fairly possessed this knowledge. He needed some detailed information of higher techniques of correlation including multiple correlation, and multiple regression equations. He studied these to his requirement. He knew nothing of the techniques of Factor Analysis. He had to apply this technique to his test. So he tried to acquire a working knowledge of this.

---

1 Op.Cit., p. V.

The investigator took about a year in doing this critical study. Then with this background, he started constructing his test.

The first step in aptitude test construction, is job analysis. We shall discuss this now.

JOB ANALYSIS

The collection of a detailed information of a profession is a prerequisite for constructing an aptitude test for that profession. This information may be of two types: (1) the types of activities a person has to take up in the profession and (2) the types of qualities in a person needed to carry out such activities successfully. This process, we would call job analysis. The first type of information is also known separately as job description.

The first question that may arise, is this: 'Why should we do job analysis at all?' Or 'How will this information be useful?'

The answer is:

Knowledge about the job is indispensable as a source of insight into functions required in the job and consequently as a basis for selection or invention of tests designed to measure those functions. It is also essential for selecting or developing criterion measures of proficiency in the job.

As it is said above, job description is primarily the description of what the worker does and the conditions under which he does it. The description of what the worker does should not only be comprehensive; it should also be specific. It should describe the major activities of a job exactly and in detail. The description of the job should be organised in such a way as to facilitate identification of those activities and characteristics that represent critical requirements of the job, as distinct from those that are relatively incidental features of it. Moreover, a job description should provide an accurate picture of significant factors in the physical, social and psychological environment in which the work must be carried out.

We have made a thorough description of a secondary school teacher's job under the heading: "Widened horizon of the duties and responsibilities of the secondary school teacher" in chapter II.

We shall discuss, therefore, the second type - the job analysis - here. By job analysis, we mean here, analysing each activity in the job to search out qualities or psychological traits which are required for carrying out the activity successfully.

Thorndike,[1] R. L. says,

---

1 Thorndike, R.L., "Personnel Selection", Test and Measurement Techniques, John Wiley & Sons, Inc., New York, 1949, p.16.

> There are two major problems in translating a job description into a good job analysis. One is to have a sound set of categories in terms of which to describe qualities of behaviour. The other is to show sagacity in identifying those categories in the job description.

What qualities are needed in a set of categories used to describe the attributes needed in a job? These are:

(1) The set of categories should be comprehensive.

(2) The set of categories should be organised and systematic.

(3) As far as possible, the categories should be independent.

(4) The categories should be psychologically meaningful.

(5) The categories should be of such a nature as to suggest testing operations for their measurement.

We can list the job analysis categories for our purpose as follows:

(1) Physical requirements: The teacher should be well-built, sound in health and without any physical defect.

(2) Sensory requirements: Each of his special senses must be sharp. Especially his eyesight should

be very bright and auditory organs must be very sharp.

(3) Intellectual requirements: A teacher should possess at least normal intelligence for doing his work successfully.

(4) Academic requirements: A teacher should be sound in subject matter. He should have good reading aptitude and possess adequate general knowledge. He should also be well-equipped with the necessary professional information.

(5) Social requirements: A teacher should have pleasing manners and good appearance. He should have willingness and capacity to understand others' behaviours, and also tact and deftness in dealing with others.

(6) Interest requirement : The teacher should have interest in profession, in reading, in child-welfare and in community at large.

(7) Emotional requirements: The teacher should be able to work under pressures of work, speed, complexity etc. He must also be emotionally stable and have good ability of personal adjustment.

PROCEDURES FOR COLLECTING JOB ANALYSIS INFORMATION

Information about a job can be obtained from a number of sources. Thorndike,[1] R. L. lists the following sources:

(1) Previous studies of the job.

(2) Analysis of documentary materials.

(3) Interviews with and interrogations of personnel.

(4) Direct experience by the job analyst.

(5) Statistical analysis of test validities.

We, for teacher's job analysis, have tapped the following sources:

(1) Reports of the studies on the same topic by other investigators.

(2) Interviews with Training College teachers.

(3) Interviews with Education Department personnel.

(4) Interviews with Secondary School principals.

(5) Interviews with veteran teachers.

(6) Interviews with successful and unsuccessful teachers.

---

1 Thorndike, R. L., Op.Cit., p. 18.

(7) Direct experience of the test constructor.

The investigator discussed with all these persons the following important points:

(1) The important work of a teacher is to carry out academic duties. He has to teach in a class and give to the children a thorough knowledge of the subject he is dealing with. What qualities in a teacher will help him discharge this duty effectively and successfully?

(2) The teachers have to see that the children build up good moral character. What types of teachers will be able to do this?

(3) A very large percentage of pupils will leave studies at the end of S.S.C.Examination and enter public life as citizens. Ours is a young democracy. We want to see that our democracy becomes successful. For this, the citizens have to discharge some duties and responsibilities. Thus, one of the manifold tasks of the secondary schools is to train the children for good citizenship. What qualities in a teacher, should there be, so as to enable him to train his students for good citizenship?

(4) Our's is a land of villages. In a village, the

village school is the centre of all activities. The teacher in a school is to organise all the activities. The teacher can carry out these activities successfully only if he has interest in people and love for the community. He has to create enthusiasm in the people and good faith in their bright future. He has to educate them and see that traditional social evils are wiped out and people take to social reforms. Is a school teacher expected to do all these? If so, what qualities are necessary in him to perform these duties?

(5) One of the aims of education, is to promote nationalism and through it, internationalism, What types of teachers can achieve this aim?

(6) The teacher has to transmit national culture in the children. He should see that the children do not fall prey to some evils such as untouchability, communalism, provincialism, religious or language fanaticism etc. He should not possess any of these evils. How to find such teachers as are above all these evils?

(7) The teacher has to do a number of things other than the academic duties in the school. He has to organise and take part in co-curricular

activities. He has to come in close contact with his colleagues and students. For this, he should have a certain type of personality. What are the traits of such personality? When can he maintain best relations with his colleagues as well as with his students?

These and many other points were discussed with about a hundred persons working in different capacities in the teaching profession. It is of no use discussing all these points here. But it must be stated that this job analysis became very interesting and helpful also. A number of factors that are likely to condition success as a secondary school teacher in Indian conditions were dug out through this process of job analysis. The investigator's eight years' experience as a secondary school teacher and about five years' experience in a Secondary Teachers' Training College helped him very much in carrying out job analysis and in understanding these factors related to teaching success.

We shall conclude the discussion of the job analysis by quoting the President[1] of a University. He says:

> The great teacher brings to his business accurate and wide knowledge, an informed technique, intelligence, energy,

---

1 Glenn Frank, quoted in "Teaching as a Career", Bulletin of the University of Wisconsin (Series 2076, 1935), published by the Bureau of Guidance and Records of the University.

> initiative, adaptability, common sense, high standards of personal character and professional achievement, singleness of purpose, sympathy, a rich social background, and a convincing sincerity of personality.

And George Herbert Palmer[1] also says, "An ideal teacher is imaginative, intellectually wealthy, stimulating and disinterest -ed."

CONSTRUCTION OF THE TEST
- SELECTION OF SUB-TESTS

The following factors which seem to contribute success in teaching, were selected tentatively:

(1) Knowledge of subject-matter.

(2) Intelligence - mental ability.

(3) General knowledge.

(4) Attitude towards community.

(5) Attitude towards children

(6) Adaptability.

(7) Professional information.

(8) Industry

(9) Sincerity

(10) Attitude towards life.

---

1 George Herbet Palmer, "The Teacher", Houghton Mifflin Company, Boston, 1908, pp. 3-30.

(11) Interest in reading.

(12) Interest in profession.

(13) Health.

(14) Personal appearance.

(15) Personality.

(16) Character.

(17) Teacher-Pupil relationship.

A cyclostyled list of these factors was prepared and circulated among the different types of personnel in the teaching profession for ranking. The sample selected for ranking the factors consisted of the following personnel.:

1.	Training College Teachers ---	20
2.	Inspectors of Secondary Schools ---	15
3.	Secondary School Principals ---	45
4.	Veteran teachers ---	20
	Total	100

Everyone was requested to rank the factors in order of their importance in contributing to the success in the secondary school teaching profession. They were to rank the most important factor/factors as 1 (One) and others in order of their importance as 2, 3, 4 and so on in rank order.

For instructions to the raters, please refer to

Appendix, A, at the end of this treatise.

The average rank calculated for each of these factors arranged in order is as follows:-

Sr. No.	Factors	Average Rank
1	Knowledge of subject-matter	1.74
2	Intelligence - mental ability	1.80
3	Attitude towards children	2.26
4	Adaptability	2.56
5	Professional information	4.10
6	Interest in profession	4.84
7	Industry	5.00
8	Sincerity	5.08
9	Teacher-pupil relationship	6.10
10	Personality	7.37
11	Character	7.37
12	General knowledge	8.19
13	Health	10.22
14	Interest in reading	14.00
15	Attitude towards community	14.05
16	Personal appearance	14.47
17	Attitude towards life	14.79

The lower the rank assigned to a factor, the greater is the contribution of that factor to a successful career in

the teaching profession. The above table indicates the order of importance given to various factors by raters for successful standing in the teaching profession. If we compare the average ranks assigned to different factors by different raters, we find that the first eight factors in the whole list are of utmost importance as they are placed in the first five ranks. These factors compare very nicely with the factors obtained from the content analysis of the existing tests of aptitude for teachers and tests for predicting teaching success. Only the seventh and eighth factors - industry and sincerity - are not included in the lists of most essential factors, that condition success in teaching profession, given by a number of investigators.

Except a few in the list of seventeen factors, the rest centre around the first six factors. We can very roughly relate the remaining factors to any one of these six factors in the following way.:

Sr. No.	Prime factor selected	Serial Number in the list of factor or factors roughly related
1	Knowledge of subject-matter	Nil
2	Intelligence - mental ability	Nil
3	Attitude towards children	9
4	Adaptability	15,17
5	Professional information	12,14
6	Interest in profession	7, 8

The remaining four factors - nos. 10, 11, 13 and 16 - do not relate to any of these six factors, so they will form their own group.

Thus our test-battery should consist of seven different sub-tests. But it was argued out that a University graduate should possess a sound knowledge of the subject-matter at least of the S.S.C. level. There are a number of subjects taught in secondary schools. Different persons will offer different subjects to teach. So it becomes almost impossible to prepare sub-tests in all these subjects. Ultimately it was decided that no separate test for subject matter need be constructed. But the prospective teachers' performances at college examinations and at secondary school examinations should be taken into consideration at the time of his selection, as a teacher, based on the results of the present aptitude test.

A prospective teacher must be physically very sound. The person must be medically examined thoroughly and if the health report is quite satisfactory, he should be thought of fit for taking up in the profession. Thus no paper-and-pencil test is needed.

After scrutinising a prospective teacher's position as regards knowledge of the subject-matter and knowing his level of aptitude for teaching, he should be interviewed thoroughly by competent judges and an attempt should be made to judge his character, personality and personal appearance. Thus

there is no need of separate paper-and-pencil test to assess the prospective teacher's personality or character.

Finally, it was agreed upon by the expert advisers and the investigator, that the present 'aptitude test for secondary school teachers' should include the following five sub-tests:

(1) Mental ability (Intelligence).

(2) Attitude towards children.

(3) Adaptability.

(4) Professional information.

(5) Interest in profession.

A number of 'aptitude tests' for teachers or tests for 'The Prediction of Teaching Success' include the above five factors.

Before we start discussing 'Writing the test items', we shall discuss the important behaviours related to these factors. This discussion will be helpful in deciding the nature of the items to be included in each sub-test.

(1) Mental Ability: Sandiford[1] and his colleagues arrived at the conclusion that ability to teach is not closely

---

1 Peter Sandiford and others, "Forecasting Teaching Ability, Bulletin No. 8, Department of Educational Research, University of Toronto, 1937, p. 60.

related to intelligence above that necessary for college graduation. In other words, if a student had sufficient intelligence to complete his University course successfully higher intelligence did not seem to be necessary of value in teaching. This sub-test, therefore, is designed to measure the normal intelligence of teachers. It includes items on word analogy, number series and reasoning ability.

(2) Attitude towards children: The teacher's main job is to work with the children. He has to impart knowledge to the children, help them build up their character and personality and see that they become good citizens. This is possible only if he can win their love and respect. If the teacher has deep understanding of their needs and problems, if he has love for them and takes interest in them, he will surely win their love and respect. If he can establish cordial relationship with them, he can be a very successful teacher. Thus, through the items of this sub-test, it is tried to reveal whether the prospective teacher will be able to establish good relations with children. A number of statements regarding pupils' behaviour are given and the testee's opinion for each is sought for.

(3) Adaptability: The teacher's main work is to teach in a class-room and thus to come in close contact with the children. He has to organise and take part in co-curricular activities also. Here, also, he has to come in close contact with the children, but in quite a different capacity and

different atmosphere. He has to adapt himself to this new situation. Moreover, he has to maintain very cordial relations with his superiors and with his colleagues. He will come in contact with the children's parents. He has to maintain good parent-teacher relationship. The society expects himself to be a good social worker. Thus he will come in close contact with people having different personality in the community. If he has no favourable attitude towards the community or has no optimistic attitude towards life, he will perhaps be unable to adjust himself with different people in the community and will not, therefore, be able to discharge certain duties of a good teacher. Through this sub-test, it is tried to measure the adaptability of the prospective teacher. A number of items, based on imaginary situations are constructed and the testee is required to answer the items to show how he will adjust himself to the situation.

(4) Professional information: The aptitude for teaching of a candidate is revealed when the classmates come to him for help. If he enjoys explaining difficult points; if he succeeds in finding ways of clarifying obscurities so that his fellows increasingly turn to him, that fact alone suggests that he probably has some aptitude for teaching. As in other callings, certainly regarding one's interest in and aptitude for teaching is enhanced by even a modest opportunity to participate in it. Teachers help those students who are seriously thinking of fitting themselves to teach, not only by

telling them more about the things teachers do, but also by giving them a chance to share in some of them - helping a backward classmate; taking charge when the teacher is called from the class-room; marking papers, preparing a laboratory demonstration, organising a nature-study trip. Opportunities like these should, of course, have come to every promising student not infrequently throughout his schooling. A good number of test-items regarding these teaching experiences may remove some of the uncertainty as to his aptitudes. Moreover, a candidate, who has aptitude for teaching knows what are the prospects in the profession, has a good deal of information of child-development and is enriched with basic information of the profession.

(5) Interest in profession: Possessed with essential potentialities, a person can develop interest in any profession. A person, who is deeply interested in the teaching profession reads books, magazines etc. on education. In case, he cannot take up the teaching profession he will select some other profession wherein he has to perform fairly similar duties. He shows interest also in the discussion of educational problems and in experiments of educational importance. To extract all these, a number of items are included in this sub-test.

WRITING THE TEST ITEMS

The mother-tongue in different states in India being different, this test could not be constructed in a particular regional language. The test was not constructed in Hindi,

though, it is a national language of India, as it is not thoroughly understood by persons in non-Hindi speaking regions. English is thoroughly understood by all college students in the country. This test was, therefore, constructed in English.

The author of this test had to construct all the test items anew for the following reasons:

(1) No aptitude test for teachers is constructed in this country.

(2) As noted earlier, the test constructor could not get any foreign aptitude test for secondary school teachers in this country. He could get only an 'Aptitude Test for Elementary School Teachers-in-Training'. Items in this test are not of much use as the present test is for secondary school teachers.

(3) Even if the investigator were able to procure some good foreign aptitude tests for teachers he would not have been able to use the items because of language intricacy and environmental influences. The items should be rewritten in simplified language and also to suit the environment here. The items in this test, therefore, are written more or less 'de novo'.

The following procedure was followed in constructing

items for each sub-test:

(1) A number of items were written.

(2) These were thoroughly discussed with his guiding professor and other experts for content validity.

(3) Correct answer to each item was decided upon through the discussion.

(4) A very small group of prospective teachers was requested to answer the items to see whether they understood the items and gave the same meanings to them and also to see whether according to them, the correct answers given by a very big majority of them, tallied with the correct answers decided by the group of experts.

(5) If there were items to which different testees gave different meanings, those items were reconstructed or discarded. If the correct answers to some items judged by the majority of testees differed from that decided by the experts, either the items were discarded or the answers were rediscussed and refixed.

(6) After this thorough scrutiny of the items in a sub-test, the items were edited for the first tryout.

(7) It was also ascertained whether the instructions given were sufficient enough to take the test and whether the instructions given at the top of each sub-test were self-explanatory.

DESCRIPTION OF ITEMS IN EACH SUB-TEST

SUB-TEST I - MENTAL ABILITY

Intelligence tests include as many as fifteen to twenty types of test items. Here, in this sub-test, we have constructed the following three types of items:

(1) Word analogy.
(2) Number series.
& (3) Reasoning ability.

In the beginning, the first type included as many as 20 items, the second type also 20 items and the third type 15 items. After the double screening as discussed in paragraph on procedure of writing test items, each type included the following number of items:

(1) Word analogy	13
(2) Number series	15
(3) Reasoning ability	8
Total:	36

SUB-TEST II - ATTITUDE TOWARDS CHILDREN

Items in this sub-test are constructed to know the prospective teacher's opinions on:

(1) Teacher-pupil relationship in general,

(2) How pupils should behave in general,

(3) How pupils should behave with teachers,

(4) How teachers should treat pupils in certain situations,

(5) General nature of children,

(6) Certain class-room techniques adopted by teachers while dealing with children behaviour problems.

To start with as many as a hundred items were constructed. After scrutiny, in all, 65 items were retained as under:

Type	Items
1	10
2	4
3	6
4	19
5	12
6	14
Total	65

SUB-TEST III - ADAPTABILITY

A secondary school teacher is likely to be put in different types of situations. A good teacher should be able

to adjust well to all these situations. These situations may arise out of the following types of relationships:

(1) Teacher-teacher relationship.

(2) Teacher-headmaster relationship.

(3) Teacher-parents relationship.

(4) Teacher-community relationship.

(5) Teacher-management relationship.

(6) Teacher's ability to adjust to situations in class-room, on playground, in community etc.

About 50 items were constructed to cover up all the six types of situations mentioned above. There was a difference of opinion among the experts regarding the usefulness of certain items in predicting teacher's ability to adapt to situations imagined in the items. These items were discarded. Finally, 28 items were selected. These items are distributed among the six categories of relationship shown above, as follows:

Type	Items
1	8
2	4
3	2
4	6
5	2
6	6
Total	28

SUB-TEST IV - PROFESSIONAL INFORMATION

What information should a prospective teacher have collected during his school or college career, if he really possessed an aptitude for teaching in schools? The investigator discussed this question with experts and the final conclusion was that the prospective teacher must possess the following types of information regarding the teaching profession:

(1) Aims of secondary school education.
(2) Child development and child's needs.
(3) Curricular and co-curricular activities.
(4) Problems of secondary school education.
(5) Learning processes and study habits.
(6) General information of teaching profession.
(7) Secondary school administration.

For each type of information, a number of items were constructed. About 75 items in total were constructed. Each item was scrutinised and ultimately 44 items were finally selected. Items on all the seven types of information are distributed as follows:

Type	No.of items
1	6
2	7
3	9
4	6
5	8
6	3
7	5
Total	44

SUB-TEST V - INTEREST IN PROFESSION

Items in this sub-test are constructed in such a way as the answer to each item will reveal whether the person who takes the test has genuine interest in the teaching profession.

This sub-test includes 10 items in all.

The items in each sub-test were arranged roughly according to difficulty level. This rough estimate of the difficulty level of items for each sub-test was obtained from the results of very small sample that was selected for administering the sub-tests on trial basis.

Each sub-test had been cyclostyled on separate sheets of paper. The testees were to write their answers on the separate answersheets.

The instructions to be given to the testees, were written at the top of each sub-test. In Appendix, B, the whole tryout test form is reproduced.

The final pattern of items for the various sub-tests is given in the table below:

Sr. No.	Name of the test	Total number of items in the test
1	Mental ability	36
2	Attitude towards children	65
3	Adaptability	28
4	Professional information	44
5	Interest in profession	10
	Total	183

CRITERION

The validity of a test is the closeness of agreement between the scores and some other objective measure of that which the test is used to measure. This other measure is called criterion. The coefficient of validity of a test is the coefficient of correlation between test scores and criterion scores.

The fact that tests of vocational aptitude so frequently have low coefficients of validity is not due alone to the difficulty of devising reliable tests of the aptitudes it is desired to measure. Adequate and reliable criteria of success against which to validate the tests are no less difficult to define and to measure.

To establish validity, then, it is first essential that an independent criterion be found.

Thorndike, R. L.[1] says,

> The most fundamental and most difficult problem in any selection research programme is to obtain satisfactory criterion measures of performance on the job, against which to validate selection procedures.

Thus a crucial problem in validating instruments of research is that of obtaining satisfactory criterion measures. The aim is to obtain a criterion measure that is statistically reliable and adequate for the purpose. The value of a criterion measure depends upon the degree to which it meets the criteria of reliability, adequacy and discrimination.

If the criterion is unreliable, the obtained correlation coefficient between predictor and criterion except for chance fluctuations will be lower than the intrinsic or "true" validity coefficient.Unreliability results from limited variability of criterion measurements and large errors of measurement. When either defect is present, measurements fail to differentiate sufficiently among the persons measured.

Inadequacy of the criterion is much more stubborn problem than unreliability. Unreliability can usually be reduced by careful editing of items, item analysis,lengthening

---

1 Thorndike, R. L., 'Personnel Selection' Test and Measurement Techniques, New York, John Wiley & Sons, Inc, London, Chapman & Hall, Ltd., 1949, p.119.

the scale, and other techniques. Jenkins[1] believes that in order to be optimally useful the criterion must be a summary measure indicating the sort of proficiency that the investigator is trying to predict. To argue the contention that the criterion itself must be "valid" seems almost to belabour the obvious.

If one asks a question, "Which is worse, the criterion or the test?" The answer is that the criterion may often be less satisfactory than the test.

We shall again quote Thorndike[2] to describe the characteristics of a satisfactory criterion. He says,

> A criterion measure is relevant as far as the knowledges, skills, and basic aptitudes required for success on it are the same as those required for performance of the ultimate task.

The present test is designed to measure an aptitude for teaching. It is essential that a satisfactory criterion of success or failure in secondary school teaching be established. The aptitude score is to be validated against this criterion score. Aptitude tests are most generally validated against external criteria. It is, therefore, of primary importance to select some satisfactory external criterion to validate our test against it.

---

1 John G. Jenkins, 'Validity for What?' Journal of Consulting Psychology, 1946, 10: 93-98.

2 Thorndike, R. L., Op.Cit., p.125.

COMMONLY USED "OUTSIDE CRITERIA"

Among the criteria frequently[1] used in validating measuring instruments are the following:

(1) the outcome of an activity - such as failure or success in school or in vocational situations;

(2) another measurement possessing known or assumed validity;

(3) associates' ratings;

(4) Self-ratings;

(5) factors isolated by factor-analysis techniques; and

(6) responses of selected groups such as inmates in an institution or members of vocational groups.

We are concerned here, with a criterion of teaching success. A number of investigators in their studies on this topic used different criteria. We shall briefly summarise these.

A criterion of teaching success is a basis used in judging the quality of teaching. Since teaching success is conditioned by many teacher traits and acquirements, the

---

1 Hsu, E.H., "A note on and some suggested methods for the determination of the validity coefficient". The Journal of Educational Psychology, 1948, 37: 305-309.

criterion is necessarily complex. The criterion most frequent-ly used in studies concerned with the prediction of teaching success, is the judgment of persons in the field, such as principals, supervisors, educational inspectors and teaching colleagues. Such judgments may involve the use of rating devices.

Other criteria include:

(1) Practice teaching marks;

(2) pupil achievement;

(3) pupil evaluation of teaching; and

(4) pupil growth.

1. Authorities are in quite general agreement that the judgments of experts is the best available criterion of teaching success.

One of the most careful studies which extensively employed the judgment of experts was conducted by Sandiford and others[1] at the University of Toronto. Using the Spearman-Brown formula, they obtained reliability coefficients of 0.888 and 0.929 respectively for the average ratings of two groups of experts and of 0.945 and 0.899 respectively for two groups of other judges. Correlations between the ratings were 0.748

1 Sandiford, Peter and others. Forecasting Teaching Ability, University of Toronto, Department of Educational Research Bulletin, No. 8, 1937, p.93.

and 0.707 respectively, indicating validities satisfactory to the investigators. Thus the criterion used by Sandiford and others was 'practice teaching marks' - the average of all lessons taught during the year - and lessons assessed by expert judges. The assumption was made that those who succeeded in practice teaching would also succeed later as regular teachers.

Henry Bowers[1] also used 'practice-teaching marks' as a criterion to validate ATEST scores against it.

2. Pupil achievement would seem to be a justifiable criterion of teaching success. However, its use is accompanied by numerous difficulties, not the least of which is an answer to the question, "what achievement?" Answers are varied and include such items as information and knowledge, attitudes, appreciations, and skills. Further difficulties are created by the lack of agreement upon what information, what knowledge and so on; the absence of valid and reliable instruments for measuring specified achievements; the possibility that pupil achievement as ordinarily measured is nearly valueless because it may be merely a measure of the efficiency with which the pupil retained factual information long enough to pass the tests; the lack of compatibility between some of the measuring instruments and the recognised objectives of education; and the

---

1 Bowers, Henry, Manual Description of the Aptitude Test for Elementary School Teachers-in-Training. J. M. Dent & Sons (Canada) Limited, Toronto Vancouver, 1948, p. 47.

rather well supported suspicion that the pupils' gain in, at least, information and knowledge is due more to his inherent ability and his habits of study than to the instruction offered by a teacher.

The studies which have employed the pupil achievement criterion have not produced high correlations between pupil gains as measured by achievement tests and independent evaluations of teachers. A positive correlation apparently exists between pupil achievement and the achievement and ability of the teacher.

3. Pupil evaluation of teaching: The criterion of pupil evaluation of teaching effectiveness has received attention from time to time. In a recent study of the personality traits of the effective teacher, Witty[1] analysed approximately 12,000 letters of pupils in grades 2-12 and found the four most frequently mentioned factors in describing the teacher who had helped them most to be co-operative, democratic attitude, kindness, and consideration for the individual, patience, and wide interest.

4. The use of pupil growth as the basis of measuring teacher efficiency is contingent upon the availability of valid instruments for measuring the factors of this growth, especially the major ones reflected in our concept of the purpose of

---

1 Witty, Paul, "An Analysis of the Personality Traits of the Effective Teacher", Journal of Educational Research, 40: 662-71; 1947.

education. A battery of achievement tests is obviously inadequate. Furthermore, teaching is only one among many factors operating to produce changes in pupil growth and achievement; the influence of factors other than teaching efficiency must be held constant in the study of growth and achievement scores. But this task of holding 'other factors' constant is practically impossible with limited resources and to that extent, the criterion of 'pupil growth' is not practical.

We shall end this discussion by quoting Sanford and Trump.[1] They say,

> Much attention is now being focussed upon the determination of a reliable criterion for judging teaching. There are at least two difficulties which hinder this determination,.....The first of these is the necessity for judging teaching in terms of a number of recognised objectives of education. This means that any acceptable criterion must recognise adequately each objective. A second difficulty is that the elements which contribute to teaching success in different communities may vary in no small measure because of the dominant values accepted and the environmental conditions operating in the communities. After a satisfactory criterion is available, it will be necessary to ascertain the factors which apparently affect a situation in such a way that good or poor teaching occurs.

---

1 Sanford, C. W. and Trump, J. L., Teacher Education - IV, Preservice Selection, Encyclopedia of Educational Research. The MacMillan Company, New York, 1956, pp. 1391-1392.

THE CRITERION SELECTED FOR THE PRESENT STUDY

The criterion finally selected for measurement of an aptitude for teaching in secondary schools was the combined marks obtained by the testees in their University B.Ed.examination in Part I and Part II examinations. Part I examination marks include the marks obtained on theory examination and marks obtained on internal marking based on various aspects such as, periodical tests, discussion during tutorial periods, library work and performance of the testee during comprehensive interviews. (It will please be noted that the scheme of internal marking followed by different training colleges is not the same. The pattern given here is followed in the Faculty of Education and Psychology, Maharaja Sayajirao University of Baroda, Baroda).

Each candidate is required to select two special methods. At the final examination he has to give two lessons one in each method. He is assessed by two examiners - one internal and one external. The internal examiner is a training college teacher and the external one is either a secondary school principal or an educational inspector or a teacher from the sister training college.

Part II marks include the marks obtained by the testees at the final examination and the average of marks obtained for all lessons taught during the year. (Again it will please be noted that the different training colleges follow different schemes. But all the schemes are naturally centred round the same nucleus).

Before deciding on this criterion, other methods discussed above, also, were considered by the investigator and the criterion was not used until its validity and reliability had been established.

The investigator thought of using any one of the other three criteria discussed above, but after deep thinking and thorough discussion with experts, it was decided that for the investigator it was not possible to use any. The use of these criteria needed very lengthy and laborious procedures which the investigator was not able to follow single-handed. Moreover, the required measuring instruments to measure pupil achievement or pupil growth are not available in this country. For pupil evaluation of teaching the experts felt that the students are not yet trained in this country, to make an objective study of teacher's work.

The validity of the selected criterion was established in the following way:

(1) Secondary school teachers' training colleges prepare the teachers to fulfil the requirements of the schools. Thus the programme of the training colleges must be in accordance with the requirements of the secondary school teaching profession. Both the training college programme and the teaching profession requirements were studied very thoroughly and found that both were

correlated to a great extent. The experts' opinion also agreed with that of the investigator.

(2) About a hundred secondary school principals were personally approached by the investigator and the following question was discussed with them:

"Do the trained teachers work more skilfully and successfully than the untrained teachers?"

The answers to the question from all the principals were identical to the extent that the trained teachers made definitely better teachers than the untrained teachers. But many of them gave the opinion that even trained teachers lacked industry and sincerity to some extent and accordingly their performance did not reach the level of efficiency, which, it should, otherwise, have.

(3) The product-moment coefficient of correlation 'r' between internal marking and total marks in University examination was calculated. It was found to be 0.716 ± 0.033 ($N = 100$). This shows that the selected criterion is much fairly valid.

After ascertaining the validity of the selected criterion, it was finally accepted for the purpose of this treatise.

SELECTION OF THE CRITERION GROUPS

The tryout form of the test was administered in March, 1957, to 153 candidates seeking admission to the Faculty of Education and Psychology, Maharaja Sayajirao University of Baroda, Baroda, as an admission test. In all, 100 candidates were admitted on the basis of scores obtained on the test and other selection criteria followed by the Faculty. From this group of 100 B.Ed. candidates of the year 1957-1958, two criterion groups were selected on the following line, for validating each item in the test.

(1) Internal marking done by different college teachers on different aspects of activities carried out by the pupil-teacher during the year, was obtained.

(2) Total marks obtained by each pupil-teacher at the University examination were obtained.

(3) The product-moment coefficient of correlation 'r' was calculated between the two.

(4) Mean, median and SD of the whole group for each type of marking were also calculated.

Below, we give a table showing individuals, their internal and University marks, etc.

TABLE NO. 3

SHOWING INDIVIDUALS FOR SELECTING CRITERION GROUPS, AND THEIR INTERNAL AND UNIVERSITY MARKS ALONGWITH THE GROUPS TO WHICH THEY ARE PLACED.

Individual	(250) Internal marks	*Rank R_1	(550) University marks	*Rank R_2	Group assigned	Remarks
1	138	40	292	44	High	-
2	116	91	256	85	Low	-
3	116	92	255	87	Low	-
4	128	64	296	37	-	Rejected
5	131	59	269	69	Low	-
6	128	65	236	92	Low	-
7	136	43	269	70	-	Rejected
8	140	32	269	71	-	Rejected
9	125	75	258	81	Low	-
10	136	44	261	79	-	Rejected
11	149	17	297	33	High	-
12	157	10	335	5	High	-
13	124	79	268	72	Low	-
14	139	37	295	39	High	-
15	149	18	320	12	High	-
16	127	69	275	63	Low	-
17	135	46	281	60	-	Rejected
18	131	60	293	43	-	Rejected
19	141	30	303	28	High	-
20	149	19	311	20	High	-

Individual	(250) Internal marks	*Rank R_1	(550) University marks	*Rank R_2	Group assigned	Remarks
21	136	45	286	55	-	Rejected
22	145	27	289	51	-	Rejected
23	144	28	284	41	High	-
24	178	2	342	2	High	-
25	140	33	279	61	-	Rejected
26	133	49	268	73	-	Rejected
27	154	14	313	18	High	-
28	148	21	315	15	High	-
29	140	34	287	53	-	Rejected
30	124	80	271	67	Low	-
31	124	81	304	25	-	Rejected
32	129	62	306	23	-	Rejected
33	133	50	301	29	-	Rejected
34	147	23	320	13	High	-
35	147	24	297	34	High	-
36	162	7	304	26	High	-
37	127	70	266	76	Low	-
38	154	15	345	1	High	-
39	105	100	215	98	Low	-
40	111	98	214	99	Low	-
41	126	73	267	75	Low	-
42	164	6	314	16	High	-
43	129	63	258	82	Low	-
44	116	93	274	64	Low	-

Individual	(250) Internal marks	*Rank R_1	(550) University marks	*Rank R_2	Group assigned	Remarks
45	167	4	290	50	-	Rejected
46	114	95	232	94	Low	-
47	166	5	335	6	High	-
48	126	74	295	40	-	Rejected
49	133	51	298	31	-	Rejected
50	149	20	296	38	High	-
51	120	87	250	90	Low	-
52	130	61	268	74	Low	-
53	123	84	274	65	Low	-
54	135	47	285	56	-	Rejected
55	147	25	314	17	High	-
56	147	26	288	52	-	Rejected
57	140	35	277	62	-	Rejected
58	133	52	256	86	-	Rejected
59	114	96	257	83	Low	-
60	127	71	270	68	Low	-
61	139	38	304	27	High	-
62	142	29	299	30	High	-
63	133	53	287	54	-	Rejected
64	127	72	266	77	Low	-
65	156	12	305	24	High	-
66	159	8	328	9	High	-
67	139	39	291	48	High	-
68	155	13	339	3	High	-

Individual	(250) Internal marks	*Rank R_1	(550) University marks	*Rank R_2	Group assigned	Remarks
69	124	82	285	57	Low	-
70	133	54	273	66	-	Rejected
71	124	83	224	97	Low	-
72	116	94	251	89	Low	-
73	109	99	237	91	Low	-
74	121	86	309	22	-	Rejected
75	120	88	229	96	Low	-
76	132	57	337	4	-	Rejected
77	119	89	266	78	Low	-
78	132	58	297	35	-	Rejected
79	128	66	260	80	Low	-
80	118	90	231	95	Low	-
81	123	85	253	88	Low	-
82	140	36	291	49	-	Rejected
83	125	76	297	36	-	Rejected
84	113	97	234	93	Low	-
85	150	16	311	21	High	-
86	133	55	322	11	-	Rejected
87	125	77	292	45	-	Rejected
88	135	48	294	42	High	-
89	128	67	257	84	Low	-
90	169	3	324	10	High	-
91	141	31	292	46	-	Rejected
92	179	1	329	7	High	-

Individual	(250) Internal marks	*Rank R_1	(550) University marks	*Rank R_2	Group assigned	Remarks
93	125	78	284	58	-	Rejected
94	148	22	329	8	High	-
95	158	9	312	19	High	-
96	137	42	292	47	High	-
97	128	68	176	100	Low	-
98	157	11	318	14	High	-
99	133	56	298	32	-	Rejected
100	138	41	282	59	-	Rejected

* It will please be noted that these ranks were not used for calculating coefficient of correlation by rank difference method but they were used in only forming the three different groups. Hence this rough method of assigning ranks was followed just to minimise labour.

University marks, as obtained from the University Office, included internal marks also. So from each individual's University marks, his internal marks were deducted and net University examination marks were obtained. It will please be remembered that our criterion includes both University total marks and internal marks also.

Now we shall proceed to calculate mean, median and SD of each type of marking.

TABLE NO. 4

DATA GROUPED FOR THE CALCULATION OF MEAN, MEDIAN AND STANDARD DEVIATION OF THE INTERNAL MARKS OF 100 INDIVIDUALS

Class interval	Mid. pt.	f	Cum f.	x'	fx'	fx'^2
170-179	174.5	2	100	4	+ 8	32
160-169	164.5	5	98	3	+15	45
150-159	154.5	9	93	2	+18	36
140-149	144.5	20	84	1	+20	20
					+61	
130-139	134.5	25	64	0	0	00
120-129	124.5	27	39	-1	-27	27
110-119	114.5	10	12	-2	-20	40
100-109	104.5	2	2	-3	- 6	18
					-53	
		N=100			$\Sigma fx'$: + 8	$\Sigma fx'^2$=218

$$C = \frac{\Sigma fx'}{N} = + \frac{8}{100} = + 0.08, \quad C^2 = 0.0064$$

$$\therefore \quad Ci = + .08 \times 10 = + 0.8$$

Mean = Assumed Mean + Ci

= 134.5 + 0.8

= <u>135.3</u>

$$\text{Median} = l + \left(\frac{\frac{N}{2} - F}{fm}\right) \times i$$

$$= 129.5 + \left(\frac{50 - 39}{25}\right) \times 10$$

= 129.5 + 4.4

= <u>133.9</u>

$$SD = \sqrt{\frac{\Sigma fx'^2}{N} - C^2} \times i$$

$$= \sqrt{\frac{218}{100} - .0064} \times 10$$

$$= \underline{14.74}$$

TABLE NO. 5

DATA GROUPED FOR THE CALCULATION OF MEAN, MEDIAN AND STANDARD DEVIATION OF THE UNIVERSITY MARKS OF 100 INDIVIDUALS

Class interval	Mid. pt.	f	Cum.f.	x'	fx'	fx'^2
330-349	339.5	6	100	+ 4	+24	96
310-329	319.5	15	94	+ 3	+45	135
290-309	299.5	29	79	+ 2	+58	116
270-289	279.5	18	50	+ 1	+18	18
					+145	
250-269	259.5	22	32	0	0	00
230-249	239.5	5	10	- 1	-5	5
210-229	219.5	4	5	- 2	-8	16
190-209	199.5	0	1	- 3	0	00
170-189	179.5	1	1	- 4	-4	16
					-17	
		N = 100			$\Sigma fx'$ = +128	$\Sigma fx'^2$ = 402

$$C = \frac{\Sigma fx'}{N} = + \frac{128}{100} = + 1.28, \quad C^2 = 1.6384$$

$$\therefore Ci = 1.28 \times 20 = 25.60$$

Mean = Assumed Mean + Ci

= 259.5 + 25.6

= 285.1

$$\text{Median} = \ell + \left(\frac{\frac{N}{2} - F}{fm} \right) \times 1$$

$$= 269.5 + \left(\frac{50 - 32}{18} \right) \times 20$$

$$= 269.5 + 20$$

$$= \underline{289.5}$$

$$SD = \sqrt{\frac{\Sigma fx'^2}{N} - C^2} \times i$$

$$= \sqrt{\frac{402}{100} - 1.6384} \times 20$$

$$= \underline{30.86}$$

These statistics for both types of marks were calculated just to show roughly how the marks are distributed in each type and therefore, reliabilities of these statistics were not calculated.

Then the product-moment coefficient of correlation, 'r' between internal and University marks, was calculated. The scatter diagram prepared for this is given on the next page.

TABLE NO. 6

SCATTER DIAGRAM BETWEEN INTERNAL MARKS AND UNIVERSITY MARKS

Internal Marks (Y - variable)	University Marks (X - variable)									
	170-189	190-209	210-229	230-249	250-269	270-289	290-309	310-329	330-349	fy
170-179	-	-	-	-	-	-	-	1	1	2
160-169	-	-	-	-	-	-	2	2	1	5
150-159	-	-	-	-	-	-	1	5	3	9
140-149	-	-	-	-	1	5	8	6	-	20
130-139	-	-	-	-	6	6	11	1	1	25
120-129	1	-	2	1	10	6	7	-	-	27
110-119	-	-	1	3	5	1	-	-	-	10
100-109	-	-	1	1	-	-	-	-	-	2
fx	1	0	4	5	22	18	29	15	6	100

The product-moment coefficient of correlation, 'r', as calculated from the above table, is 0.716 ± 0.033.

$$P.E._r = .6745 \times \frac{1 - r^2}{\sqrt{N}}$$

$$= .6745 \times \frac{1 - (.716)^2}{\sqrt{100}}$$

$$= 0.033$$

After all such possible information about the individuals was collected, they were divided into three groups - high - middle - low - in the following way:

Group I - High group - 33 individuals.

Pupil-teachers, getting HIGH score on internal marking and HIGH on University examination, were placed in this group.

Group II - Middle group - 34 individuals.

Pupil-teachers, getting HIGH score on internal marking and LOW score on University examination or vice-versa, were placed in this group.

Group III - Low group - 33 individuals.

Pupil-teachers, getting LOW score on internal marking and LOW score on University examination, were placed in this group.

Groups I and III were tentatively selected as criteria

-n

groups and group II was discarded. In table No. 3, against each individual, his group is mentioned.

Group I - High group - Total: 33.

1, 11, 12, 14, 15, 19, 20, 23, 24, 27, 28, 34, 35, 36, 38, 42, 47, 50, 55, 61, 62, 65, 66, 67, 68, 85, 88, 90, 92, 94, 95, 96, 98.

Group III - Low group - Total: 33.

2, 3, 5, 6, 9, 13, 16, 30, 37, 39, 40, 41, 43, 44, 46, 51, 52, 53, 59, 60, 64, 69, 71, 72, 73, 75, 77, 79, 80, 81, 84, 89, 97.

For further scrutiny of the two criterion groups, the school principals, where the teachers belonging to these groups were serving after their training, were requested to give their estimation of the teachers' work after training. They were requested to rate the teachers on a seven-point scale - A, B, C^{+}, C, C^{-}, D, E. If, in his opinion, the teacher was the best, 'A' grade was to be given; while if a teacher was found unaccept-able to the profession, he was to be given 'E' grade, and so on. (Please refer to Appendix C).

The teacher about whom the principal's opinion was much different from the forecast of the University marks, was discarded from the group. Thus much more weightage was given to the principal's estimation of the teacher's work. This should not be objectionable as ultimately the principal is a

right judge of the teacher's work, as he is observing him and his work constantly. This is again, a test of the validity of the criterion.

In a table below, we give the principal's estimation of the teacher's work after training.

TABLE NO. 7

PRINCIPAL'S ESTIMATION OF THE TRAINED TEACHERS

High Group			Low Group		
Individual	Principal's estimation	Remarks	Individual	Principal's estimation	Remarks
1	C^+	-	2	D	-
11	B	-	3	D	-
12	A	-	5	C^-	-
14	B	-	6	C^-	-
15	A	-	9	C^-	--
19	C^+	-	13	C^-	-
20	B	-	16	C^-	-
23	C	Rejected	30	D	-
24	A	-	37	C^-	-
27	B	-	39	D	-
28	C^+	-	40	D	-
34	B	-	41	C^-	-
35	C^+	-	43	C^-	-
36	C^+	-	44	C^-	-
38	A	-	46	C	Rejected

High Group			Low Group		
Individual	Principal's estimation	Remarks	Individual	Principal's estimation	Remarks
42	B	-	51	D	-
47	B	-	52	D	-
50	C^+	-	53	C	Rejected
55	C	*	59	D	-
61	C^*	-	60	D	-
62	C	*	64	C^+	Rejected
65	B	-	69	D	-
66	B	-	71	D	-
67	C^-	Rejected	72	D	-
68	B	-	73	D	-
85	C	*	75	D	-
88	C	Rejected	77	C^-	-
90	A	-	79	C^-	-
92	A	-	80	C^+	Rejected
94	B	-	81	D	-
95	C^+	-	84	D	-
96	C^-	Rejected	89	C^-	-
98	C^+	-	97	D	-

From the above table, we can see that from 'low group' we have rejected individuals getting 'C' or 'C^+' , and from 'high group', we have rejected individuals getting 'C' or 'C^-'. From 'high group', we have not rejected individuals, 55, 62 and

85 though they got 'C', just to maintain equal number of individuals in both the groups. These three individuals were retained because their internal and University marks correlate much better than those of the rejected individuals with 'C' grade.

The product-moment coefficient of correlation 'r', between University marks and principal's estimation was calculat-ed. The correlation-table prepared for the same, is given below:

TABLE NO. 8

SCATTER-DIAGRAM BETWEEN UNIVERSITY MARKS AND PRINCIPALS' ESTIMATION

University marks (Y - variable)	Principals' Estimation (X - variable)							
	E	D	C^-	C	C^+	B	A	fy
330-349	-	-	-	-	-	2	3	5
310-329	-	-	-	2	3	6	3	14
290-309	-	-	2	3	6	3	-	14
270-289	-	3	2	1	-	-	-	6
250-269	-	7	9	-	1	-	--	17
230-249	-	2	1	1	1	-	-	5
210-229	-	4	-	-	-	-	-	4
190-209	-	-	-	-	-	-	-	0
170-189	-	1	-	-	-	-	-	1
fx	0	17	14	7	11	11	6	66

The product-moment coefficient of correlation, 'r',

as calculated from the above table, is 0.803 ± 0.0295.

$$P.E._r = .6745 \times \frac{1 - r^2}{\sqrt{N}}$$

$$= .6745 \times \frac{1 - (.803)^2}{\sqrt{66}}$$

$$= 0.0295$$

This shows that the University marks and principals' estimation correlate sufficiently high.

On the basis of principals' estimation four individuals from each of the two groups were rejected.

The two criterion groups were, then, finally selected. Each group includes 29 individuals. The two groups finally formed were as follows:

A. HIGH GROUP: (N = 29)

1, 11, 12, 14, 15, 19, 20, 24, 27, 28, 34, 35, 36, 38, 42, 47, 50, 55, 61, 62, 65, 66, 68, 85, 90, 92, 94, 95, 98.

B. LOW GROUP: (N = 29)

2, 3, 5, 6, 9, 13, 16, 30, 37, 39, 40, 41, 43, 44, 51, 52, 59, 60, 69, 71, 72, 73, 75, 77, 79, 81, 84, 89, 97.

Group A - High group - consists of individuals having high measure of aptitude for teaching and group B - Low group -

consists of individuals having low measure of aptitude for teaching or in other words, group A consists of good teachers and group B consists of comparatively poor teachers.

Further inquiries about these teachers made last year (in the year, 1960) showed that many of the teachers in group A were doing excellent work in the profession. Not less than 15 per cent of them were working as H.M.'s in rural schools and two of them were appointed as lecturers in the Training Colleges. While from group B, two teachers had left the profession to start business, and the rest were working quite ordinarily and on normal routine lines.

This information also adds to testing the validity of the criterion.

ADMINISTRATION OF THE TRYOUT-TEST

As previously stated, the tryout test was administered to 153 candidates, seeking admission to the B.Ed. class in the Faculty of Education and Psychology, Maharaja Sayajirao University of Baroda, Baroda, for the year 1957-1958. They took the test in the second week of March, 1957. The test was administered under the direct supervision of the investigator and his colleagues in the Faculty.

This time, the instructions were not written on the test forms. They were written on a separate piece of paper and given by the investigator orally at the time of administering the test. They were explained how to take the test and

were requested to attempt each item in the test and write answers to the items sincerely and honestly. Moreover, they were told that they would be given as much time as they needed and so they should read and understand each item thoroughly and take the test patiently. They were also permitted to ask the meanings of the difficult words they did not know and also to ask for any clarification regarding taking the test.

The maximum time taken by the slowest candidate in the group was two and a half hours. Thus he took approximately 3/4th of a minute to answer an item in the test. This gave a rough estimate as to 'of how much time the final test would be.' It was estimated, at this time that the final test would be of about an hour and a half.

From the experience of administering the test, the investigator found that some of the instructions, he had previously prepared, were not needed. As the sample consisted of matured people, some of the formal instructions were not at all needed. He, therefore, scored out those instructions. On the contrary, he had to add some instructions on 'how to take the test'. He found that some illustrations for each sub-test were needed. He, thus, from the experience of this test administration, revised all the instructions and wrote them again for the pilot form of the test.

SCORING

Correct answers to each item in each sub-test had

been decided earlier through discussions with experts. The investigator, therefore, could start scoring immediately after the testing was over. The scoring was done by the investigator himself. This enabled him to study the nature of the answers - responses selected - by the testees. This helped him a lot at a later stage in revising the items for the pilot form of the test.

SUB-TEST I - MENTAL ABILITY

The answers are of the type - right or wrong. One score was to be given to each correct answer given by the testee. There were in all 36 items in this test. Maximum score obtainable was, therefore, 36.

SUB-TEST II - ATTITUDE TOWARDS CHILDREN

Here, the items are to be scored on five-point scale. The most favourable answer was to be given 5 scores and the most unfavourable answer was to be given 1 score. The total score was to be divided by 5 and thus the average maximum score obtainable here was 65, as there are 65 statements in this sub-test.

SUB-TEST III - ADAPTABILITY

The items in this test are of multiple choice type. Credit was to be given only to the best choice selected. There were 28 items in this test and maximum obtainable score was, therefore, also 28.

SUB-TEST IV - PROFESSIONAL INFORMATION

The items, here, are also of multiple-choice type. Credit was to be given only to the best choice selected. There were 44 items in the test and so the maximum score obtainable was also 44.

SUB-TEST V - INTEREST IN PROFESSION

In all, there are 10 items in this test. But the nature of all items is not the same. The mode of scoring the items, therefore, cannot be the same.

The first six items are of multiple-choice type. The best choice selected was to be given one score. Total scores would be 6.

In items 7, 8, 9, 10, some sort of selections were to be made. In each item, three acceptable answers were fixed. These were in order of priority. The best one was to be scored 3, the second one, 2 and the last one, 1. The maximum score obtainable in this sub-test was, therefore, 18.

The total maximum score obtainable for the whole test was, therefore, 191.

Next step in the test construction was, naturally, of validating each item. This was not possible unless the two criterion groups were selected. Formation of criterion groups was based on internal and University marks. Marks could

only be available after University results were declared. So the investigator had to delay his work till the end of April, 1958. After getting the necessary data, the groups were formed. (In the preceding pages we have discussed the whole procedure followed for forming the groups). The principals' opinion about the working in the schools of the members of both the groups - high and low - was obtained by the middle of July, 1958, i.e. after about a month they had worked in the schools after their training. Thus, after the groups were finally decided, the item - validation process was carried out.

ITEM VALIDITY

The need to find out an index of validity for each of the items in the test, before it is preferred for inclusion in the final form of the test, is for establishing the validity of the test as a whole. If the item validity is not calculated in the beginning, one does not know how the test will turn out when it is finalised. On the contrary, if the validity of each item is determined right from the beginning and only valid items are included in the test, there can be no doubt that the test is bound to be valid.

The validity of each item in this test was ascertained by using the two criterion groups of prospective teachers selected by a very lengthy and laborious but much reliable and thorough process.

There are numerous indices and procedures for -

determining item validity.

PSYCHOMETRIC PRECISION OF ITEMS

(1) Ferguson[1] proposed that the processes of the method of constant stimuli be adapted to the study of items.

(2) Turnbull[2] has proposed a graphic procedure based on the same principles (method of constant stimuli).

Turnbull's process also yields an estimate of correlation of item with criterion.

INDICES OF DISCRIMINATION BETWEEN HIGH AND LOW GROUPS

A very common step is to divide the total sample into two groups on the basis of the criterion.

(1) Critical-ratio test: The percents of correct responses in two separate criterion groups are obtained, and the differences between these pairs of percents are computed for comparison with their own standard errors. The resulting critical-ratio for each item indicates how likely it is that a given item actually differentiates between the two criterion groups.

---

1 Ferguson, G. A., Item Selection by the Constant Process, Psychometrika, 1942, pp. 19-29.

2 Turnbull, W.W., A normalised graphic method of item analysis. Journal of Educational Psychology, 1946, pp. 129-141.

Mosier and McQuitty[1] have published an abac for graphic solution for the critical ratio in the item-analysis situation.

(2) Guilford[2] proposed chi-square as a measure of item discrimination. He has also presented an abac for determining under limited conditions whether the difference between these two pairs of percents yields a significant chi-square.

(3) Davis[3] describes a chi-test attributed to Cureton, which is especially convenient when samples are small. The chi-test involves no assumptions regarding the shape of the distribution of the traits measured by either the criterion or item scores.

For validating the items in this test we have used the chi test proposed by Cureton, when the number of testees in the high group who mark the item correctly is greater than the number in the low group who mark it correctly, the following formula[4] for computing chi, is used.

---

1 Mosier, C.I., and McQuitty, J.V., Methods of Item Validation and abacs for item-test construction and critical ratio of upper-lower differences, Psychometrika,1940, 57-65.

2 Guilford, J.P. "The Phi Coefficient and chi-square as indices of Item Validity," Psychometrika, 1941, 11-19.

3 Lindquist, E. F. (ed.),Op.Cit., p.288.

4 Lindquist, E. F. (ed.), Op.Cit., p.289.

$$Chi = \frac{R_h - R_\ell - 1}{\sqrt{R_t \left(1 - \frac{R_t}{N_t - NR_t} \right)}}$$

Where: R_h = the number of testees in the high group who mark the item correctly,

R_ℓ = the number of testees in the low group who mark the item correctly,

$R_t = R_h + R_\ell$,

N_t = the number of testees in the high and low groups,

NR_t = the number of testees in the high and low groups who do not reach the item in the time limit.

As the testees were given quite enough time to take the test, there were no testees who did not reach the item, and NR_t, therefore, is zero, and so the formula was reduced to:

$$Chi = \frac{R_h - R_\ell - 1}{\sqrt{R_t \left(1 - \frac{R_t}{N_t} \right)}}$$

If the number of testees in the high group who mark the item correctly is smaller than the number in the low group who mark it correctly, a different formula[1] is to be used. There were seven such items in our test. As we could not include

---

1 Lindquist, E. F., (ed.), Op.Cit., p.289,

such items in our test, they were discarded outright. Chi-test was, therefore, not used for such items and hence the use of this formula did not arise at all.

Following steps were followed to apply Cureton's chi test to our test items:

(1) R_h value of each item was found out.

(2) $R\ell$ value of each item was found out.

(3) $R_t = R_h + R\ell$, was calculated for each item.

(4) N_t = 58 throughout ('High' 29 + 'low' 29).

(5) Chi, for each item, was calculated.

(6) It was noted whether each chi was significant. For this abridged table[1] made by Cureton of Table 8 in Statistical Tables for Biological, Agricultural, and Medical Research was used by the investigator. (This table can only be used if the two groups selected are invariably of equal size).

The values of R_h, $R\ell$, R_t, chi etc. for each item, are shown in a table No. 9 , given below. In the table, against each item, it is mentioned whether it is retained or rejected. The value of N_t is 58 for all the 183 items.

---

1 Lindquist, E. F. (ed.), Op.Cit., p.290.

TABLE NO. 9

A TABLE SHOWING THE VALUES OF R_h, R_ℓ, R_t Chi, etc. FOR EACH ITEM. $N_t = 58$.

S = (chi) SIGNIFICANT AT 0.05 level
N.S. = (chi) NOT SIGNIFICANT AT 0.05 level

Sub-test	Item No.	R_h	R_ℓ	R_t	Chi	Whether significant or not	Whether the item is retained or rejected.
1	2	3	4	5	6	7	8
I	1	23	9	32	3.43	S	Retained
	2	19	7	26	2.90	S	Retained
	3	26	11	37	3.82	S	Retained
	4	24	10	34	3.47	S	Retained
	5	25	8	33	4.24	S	Retained
	6	17	3	20	3.59	S	Retained
	7	21	6	27	3.69	S	Retained
	8	22	9	31	3.16	S	Retained
	9	21	6	27	3.69	S	Retained
	10	-	-	-	-	-	*
	11	19	12	31	1.58	N.S.	Rejected
	12	11	5	16	1.47	N.S.	Rejected
	13	16	4	20	3.04	S	Retained
	14	27	20	47	2.01	N.S.	Rejected
	15	29	23	52	2.16	N.S.	Rejected
	16	25	14	39	2.80	S	Retained
	17	21	5	26	3.96	S	Retained
	18	20	10	30	2.37	N.S.	Rejected

Sub-test	Item No.	R_h	R_l	R_t	Chi	Whether significant or not.	Whether the item is retained or rejected
1	2	3	4	5	6	7	8
I	19	25	13	38	3.04	S	Retained
	20	19	7	26	2.90	S	Retained
	21	26	15	41	2.88	S	Retained
	22	28	15	43	3.60	S	Retained
	23	-	-	-	-	-	*
	24	28	17	45	3.15	S	Retained
	25	17	4	21	3.28	S	Retained
	26	21	8	29	3.15	S	Retained
	27	18	4	22	3.52	S	Retained
	28	23	11	34	2.93	S	Retained
	29	17	4	21	3.28	S	Retained
	30	18	7	25	2.65	S	Retained
	31	21	9	30	2.89	S	Retained
	32	16	3	19	3.36	S	Retained
	33	25	12	37	3.28	S	Retained
	34	22	7	29	3.68	S	Retained
	35	28	15	43	3.60	S	Retained
	36	24	11	35	3.22	S	Retained
II	1	25	17	42	2.06	N.S.	Rejected
	2	22	10	32	2.90	S	Retained
	3	24	11	35	3.22	S	Retained
	4	20	9	29	2.63	S	Retained
	5	28	23	51	1.61	N.S.	Rejected

Sub-test	Item No.	R_h	R_l	R_t	Chi	Whether significant or not	Whether the item is retained or rejected
1	2	3	4	5	6	7	8
II	6	26	15	41	2.88	S	Retained
	7	26	18	44	2.15	N.S.	Rejected
	8	-	-	-	-	-	*
	9	15	8	23	1.61	N.S.	Rejected
	10	21	10	31	2.63	S	Retained
	11	26	15	41	2.88	S	Retained
	12	27	14	41	3.46	S	Retained
	13	20	8	28	2.89	S	Retained
	14	15	11	26	0.79	N.S.	Rejected
	15	25	17	42	2.06	N.S.	Rejected
	16	17	5	22	2.98	S	Retained
	17	24	11	35	3.22	S	Retained
	18	14	8	22	1.35	N.S.	Rejected
	19	15	3	18	3.12	S	Retained
	20	24	12	36	2.98	S	Retained
	21	14	9	23	1.07	N.S.	Rejected
	22	25	11	36	3.52	S	Retained
	23	13	2	15	3.00	S	Retained
	24	26	13	39	3.36	S	Retained
	25	19	13	32	1.32	N.S.	Rejected
	26	17	5	22	2.98	S	Retained
	27	20	8	28	2.89	S	Retained

Sub-test	Item No.	R_h	R_l	R_t	Chi	Whether significant or not	Whether the item is retained or rejected
1	2	3	4	5	6	7	8
II	28	19	6	25	3.18	S	Retained
	29	18	7	25	2.65	S	Retained
	30	13	9	22	0.81	N.S.	Rejected
	31	20	9	29	2.63	S	Retained
	32	12	1	13	3.15	S	Retained
	33	24	11	35	3.22	S	Retained
	34	17	6	23	2.68	S	Retained
	35	20	8	28	2.89	S	Retained
	36	15	3	18	3.12	S	Retained
	37	21	7	28	3.42	S	Retained
	38	19	6	25	3.18	S	Retained
	39	15	4	19	2.80	S	Retained
	40	17	4	21	3.28	S	Retained
	41	13	9	22	0.81	N.S.	Rejected
	42	13	2	15	3.00	S	Retained
	43	25	14	39	2.80	S	Retained
	44	17	3	20	3.59	S	Retained
	45	22	11	33	2.65	S	Retained
	46	-	-	-	-	-	*
	47	12	7	19	1.12	N.S.	Rejected
	48	8	7	15	0.00	N.S.	Rejected
	49	20	6	26	3.43	S	Retained

Sub-test	Item No.	R_h	R_l	R_t	Chi	Whether significant or not	Whether the item is retained or rejected
1	2	3	4	5	6	7	8
II	50	19	6	25	3.18	S	Retained
	51	9	6	15	0.60	N.S.	Rejected
	52	23	11	34	2.93	S	Retained
	53	20	8	28	2.89	S	Retained
	54	19	6	25	3.18	S	Retained
	55	18	6	24	2.93	S	Retained
	56	26	15	41	2.88	S	Retained
	57	14	0	14	3.99	S	Retained
	58	13	5	18	1.99	N.S.	Rejected
	59	8	2	10	1.74	N.S.	Rejected
	60	27	14	41	3.46	S	Retained
	61	18	7	25	2.65	S	Retained
	62	25	14	39	2.80	S	Retained
	63	7	3	10	1.04	N.S.	Rejected
	64	4	0	4	1.02	N.S.	Rejected
	65	-	--	-	-	-	*
III	1	26	15	41	2.88	S	Retained
	2	26	15	41	2.88	S	Retained
	3	24	12	36	2.98	S	Retained
	4	19	6	25	3.18	S	Retained
	5	27	15	42	3.23	S	Retained
	6	26	19	45	1.89	N.S.	Rejected

Sub-test	Item No.	R_h	R_l	R_t	Chi	Whether significant or not	Whether the item is retained or rejected
1	2	3	4	5	6	7	8
III	7	12	0	12	3.57	S	Retained
	8	16	3	19	3.36	S	Retained
	9	17	5	22	2.98	S	Retained
	10	19	6	25	3.18	S	Retained
	11	20	9	29	2.63	S	Retained
	12	15	4	19	2.80	S	Retained
	13	25	11	36	3.52	S	Retained
	14	27	16	43	3.00	S	Retained
	15	18	7	25	2.65	S	Retained
	16	25	14	39	2.80	S	Retained
	17	28	17	45	3.15	S	Retained
	18	29	17	46	3.57	S	Retained
	19	8	6	14	0.31	N.S.	Rejected
	20	16	4	20	3.04	S	Retained
	21	29	17	46	3.57	S	Retained
	22	16	5	21	2.73	S	Retained
	23	7	3	10	1.04	N.S.	Rejected
	24	15	4	19	2.80	S	Retained
	25	28	17	45	3.15	S	Retained
	26	25	13	38	3.04	S	Retained
	27	18	7	25	2.65	S	Retained
	28	21	10	31	2.63	S	Retained

Sub-test	Item No.	R_h	R_f	R_t	Chi	Whether significant or not	Whether the item is retained or rejected
1	2	3	4	5	6	7	8
IV	1	-	-	-	-	-	*
	2	17	6	23	2.68	S	Retained
	3	16	4	20	3.04	S	Retained
	4	13	2	15	3.00	S	Retained
	5	20	9	29	2.63	S	Retained
	6	17	6	23	2.68	S	Retained
	7	5	0	5	1.87	N.S.	Rejected
	8	18	7	25	2.65	S	Retained
	9	11	0	11	3.35	S	Retained
	10	17	5	22	2.98	S	Retained
	11	3	0	3	1.19	N.S.	Rejected
	12	20	9	29	2.63	S	Retained
	13	21	7	28	3.42	S	Retained
	14	18	6	24	2.93	S	Retained
	15	20	8	28	2.89	S	Retained
	16	18	7	25	2.65	S	Retained
	17	19	7	26	2.90	S	Retained
	18	29	23	52	2.16	N.S.	Rejected
	19	11	0	11	3.35	S	Retained
	20	11	0	11	3.35	S	Retained
	21	13	2	15	3.00	S	Retained
	22	21	10	31	2.63	S	Retained
	23	20	5	25	3.71	S	Retained

Sub-test	Item No.	R_h	R_l	R_t	Chi	Whether significant or not	Whether the item is retained or rejected
1	2	3	4	5	6	7	8
IV	24	-	-	-	-	-	*
	25	20	3	23	4.29	S	Retained
	26	25	10	35	3.76	S	Retained
	27	18	6	24	2.93	S	Retained
	28	11	3	14	2.15	N.S.	Rejected
	29	16	5	21	2.73	S	Retained
	30	25	14	39	2.80	S	Retained
	31	23	12	35	2.68	S	Retained
	32	18	6	24	2.93	S	Retained
	33	25	20	45	1.26	N.S.	Rejected
	34	19	6	25	3.18	S	Retained
	35	11	0	11	3.35	S	Retained
	36	26	11	37	3.82	S	Retained
	37	28	15	43	3.60	S	Retained
	38	23	12	35	2.68	S	Retained
	39	23	11	34	2.93	S	Retained
	40	15	4	19	2.80	S	Retained
	41	3	0	3	1.19	N.S.	Rejected
	42	17	5	22	2.98	S	Retained
	43	18	7	25	2.65	S	Retained
	44	18	6	24	2.93	S	Retained
V	1	12	4	16	2.06	N.S.	Rejected

Sub-test	Item No.	R_h	R_l	R_t	Chi	Whether significant or not	Whether the item is retained or rejected
V	2	22	10	32	2.90	S	Retained
	3	29	18	47	3.35	S	Retained
	4	23	12	35	2.68	S	Retained
	5	19	7	26	2.90	S	Retained
	6	20	8	28	2.89	S	Retained
	7	19	6	25	3.18	S	Retained
	8	29	17	46	3.57	S	Retained
	9	28	17	45	3.15	S	Retained
	10	15	4	19	2.80	S	Retained

The chi test involves no assumptions regarding the shape of the distribution of the traits measured by either the criterion or item scores.

Asterisk in the last column shown in the table indicat-es that the particular item was discarded before applying the chi-test, as either it was answered correctly by all in both the groups or the number of testees in the high group who marked the item correctly was smaller than the number in the low group who marked it correctly.

It is seen in the above table that out of 183 items, as many as 40 items were rejected as these items did not show any significant discrimination between the two criterion groups.

The remaining 143 items were retained in the test.

The table below, shows the number of valid items in the different sub-tests.

TABLE NO. 10

SHOWING NUMBER OF VALID ITEMS IN DIFFERENT SUB-TESTS

Sr. No.	Sub-test	No. of items in the test
1	Mental Ability	29
2	Attitude Towards Children	44
3	Adaptability	25
4	Professional Information	36
5	Interest in Profession	9
	Total	143

Some changes were made in the construction of the items 6,7,8 & 9 in sub-test five. More than four choices were previously given in each of these items. Some of these choices did not work well. Moreover, this created some scoring problems also. The number of choices given in each of these items were, therefore, reduced to four to bring the items on the similar pattern with the other items in the test. The choices which worked better in the tryout testing were selected for the purpose.

The pilot form of the test was, thus,ready. It contain-ed 143 items in all. The whole test is reproduced in Appendix D.