CHAPTER VIII

ADMINISTRATION OF THE FINAL FORM OF THE TEST FOR STANDARDISATION

In the preceding chapter, we saw that 120 items were selected finally on the basis of experts' opinion, item validity, internal consistency indices, and item difficulty indices. All these 120 items were distributed among the five sub-tests as follows:

TABLE NO. 14

SHOWING DISTRIBUTION OF ITEMS AMONG SUB-TESTS SELECTED FOR FINAL TEST

Sub-Test number	Total number of items in the sub-test
I	26
II	35
III	20
IV	30
V	9
Total	120

A cursory glance at table No. 11 in the previous chapter, showing difficulty values of different items in

different sub-tests will reveal that the items in sub-tests III or IV are comparatively more difficult than those in the sub-tests I, II or V. In the final test form, the order of the sub-tests had, therefore, to be rearranged according to the difficulty levels of the sub-tests.

Moreover, some printing requirements also necessitated the rearrangement of the sub-tests in the final form.

The change in the order of the sub-tests is shown in the following table:

TABLE NO. 15

SHOWING THE NEW ORDER OF THE SUB-TESTS IN THE FINAL TEST

Sr.No.of the sub-test in Pilot form	Sr.No.of the sub-test in Final form	Title of the sub-test	Total no.of items in the sub-test
II	I	Attitude towards children	35
I	II	Mental ability	26
V	III	Interest in profession	9
III	IV	Adaptability	20
IV	V	Professional information	30
-	-	Total	120

Henceforth, in this treatise, the reference of the sub-tests will be made by the new serial numbers of the sub-tests.

According to the general principles of the test

construction, the items in each sub-test should have been arranged according to their difficulty values, the easiest being at the top and the hardest at the bottom. This general rule of arranging the items was not strictly followed here. The following points were taken into consideration while rearranging the items in each sub-test in the final form of the test.

(1) Difficulty value of each item,

(2) the type or group in a sub-test to which the item belonged,

(3) the choice letter of correct response for each item, and

(4) the form of the item.

The distractors attached to each multiple-choice item were transposed to provide roughly an equal number of correct responses for each choice letter. This would help probably in reducing guessing while answering, to some extent.

The rearrangement of the items in each sub-test made it imperative to change the serial numbers that were given to the items in the pilot test.

For reading the internal-consistency indices and difficulty values of items included in the final test, from table No. 11 on page No. 192, in chapter VII, it is necessary to know their corresponding serial numbers in the pilot

Table No. 16.

test. The table below fulfils this requirement:

Sr.No. of the sub-test in final form	Sr.No. of the item in final form	Sr.No. of the sub-test in pilot form	Sr.No. of the item in pilot form	Sr.No. of the sub-test in final form	Sr.No. of the item in final form	Sr.No. of the sub-test in pilot from	Sr.No. of the item in pilot form
I	1	II	5	I	23	II	33
"	2	"	21	"	24	"	38
"	3	"	27	"	25	"	2
"	4	"	14	"	26	"	30
"	5	"	36	"	27	"	9
"	6	"	42	"	28	"	16
"	7	"	44	"	29	"	40
"	8	"	25	"	30	"	39
"	98	"	35	"	31	"	19
"	10	"	13	"	32	"	15
"	11	"	23	"	33	"	32
"	12	"	24	"	34	"	26
"	13	"	8	"	35	"	41
"	14	"	10	IIA	1/1	IA	3
"	15	"	22	"	2/2	"	5
"	16	"	29	"	3/3	"	4
"	17	"	34	"	4/4	"	2
"	18	"	18	"	5/5	"	9
"	19	"	1	"	6/6	"	8
"	20	"	11	"	7/7	"	7
"	21	"	17	"	8/8	"	1
"	22	"	28	"	9/9	"	6

Sr.No. of the sub-test in final form	Sr.No. of the item in final form	Sr.No. of the sub-test in pilot form	Sr.No. of the item in pilot form	Sr.No. of the sub-test in final form	Sr.No. of the item in final form	Sr.No. of the sub-test in pilot form	Sr.No. of the item in pilot form
IIB	1/10	IB	5	III	6	V	6
"	2/11	"	3	"	7	"	7
"	3/12	"	7	"	8	"	8
"	4/13	"	1	"	9	"	9
"	5/14	"	6	IV	1	III	19
"	6/15	"	2	"	2	"	17
"	7/16	"	10	"	3	"	22
"	8/17	"	11	"	4	"	12
"	9/18	"	9	"	5	"	23
"	10/19	"	8	"	6	"	3
IIC	1(a)/20	IC	7	"	7	"	13
"	1(b)/21	"	7	"	8	"	14
"	2/22	"	3	"	9	"	5
"	3/23	"	5	"	10	"	15
"	4/24	"	6	"	11	"	24
"	5/25	"	1	"	12	"	4
"	6/26	"	4	"	13	"	10
III	1	V	2	"	14	"	25
"	2	"	1	"	15	"	7
"	3	"	3	"	16	"	2
"	4	"	5	"	17	"	9
"	5	"	4	"	18	"	11

Sr.No. of the sub-test in final form	Sr.No. of the item in final form	Sr.No. of the sub-test in pilot form	Sr.No. of the item in pilot form	Sr.No. of the sub-test in final form	Sr.No. of the item in final form	Sr.No. of the sub-test in pilot form	Sr.No. of the item in pilot form
IV	19	III	8	V	15	IV	18
"	20	"	6	"	16	"	14
V	1	IV	4	"	17	"	22
"	2	"	21	"	18	"	3
"	3	"	30	"	19	"	26
"	4	"	11	"	20	"	34
"	5	"	32	"	21	"	9
"	6	"	15	"	22	"	8
"	7	"	31	"	23	"	20
"	8	"	5	"	24	"	19
"	9	"	24	"	25	"	17
"	10	"	13	"	26	"	7
"	11	"	1	"	27	"	25
"	12	"	2	"	28	"	36
"	13	"	29	"	29	"	28
"	14	"	10	"	30	"	16

THE FINAL TEST - AN ASSEMBLY OF FIVE SUB-TESTS

The final test is, thus, an assembly of five sub-tests.

The whole battery of the sub-tests is a printed one.

It is printed in the form of an eighteen page booklet.

On the front cover page of the test booklet, seven standardised instructions to the testees are printed. The test constructor's and his guiding professor's names are printed on the second page.

The instructions as to how to answer items in each sub-test are given at the top of each sub-test together with an illustration for the same.

The remaining sixteen pages in the test-booklet are covered by the five sub-tests in the following way:

Sub-test No.	Title	Pages covered in the test-booklet
I	Attitude towards children	3 and 4
II	Mental ability	5 and 6
III	Interest in profession	7 and 8
IV	Adaptability	9,10,11 and 12
V	Professional information	13,14,15,16,17 and 18

As the testees are expected to write their answers in the separate answer-book provided for the same, they were not to write anything in the test-booklet. The same test-booklet could, therefore, be used again and again. As the investigator did not expect more than 150 testees at a time at any testing centre, 200 copies of the test-booklet were printed.

A copy of the printed test-booklet is attached at the end of this treatise as Appendix, E.

THE ANSWERBOOK

The answerbook is also a printed one. It contains four pages.

On the front page, the testees are required to supply the necessary particulars about them before they start answering the items in the test.

A table showing the separate scores in each sub-test and the total test-score obtained by the testee, is also given on the front page of the answerbook.

The testees are to give their responses to the items in first sub-test on second page, those to the second and third test items on page three and those to the fourth and fifth test items on page four of the answerbook.

At the end of each sub-test in the answerbook a line, "Please make sure that you have omitted no item in the test -___" just to remind the testees each time, they pass on to the next sub-test, that they are not to omit any test item.

The investigator had estimated that the total sample would be of about 700 testees. So one thousand copies of the answerbooks were printed.

A copy of the printed answerbook is attached at the

end of this treatise as Appendix, F.

SAMPLE

Generally Arts and Science graduates are recruited as teachers in our secondary schools. This test is designed to measure their aptitude for teaching. The test should, therefore, be naturally administered to them to collect data for standardising it. Arts and Science graduates form a very large population. It was physically impossible and also uneconomical to administer the test to the whole population. A section of this population had, then, to be selected for administering the test for collecting data for standardisation. This section of population is known as a 'sample' in the statistical terms. The question is, whether we can select any section of the population we like and treat it as our sample, for test standardisation. Naturally, we cannot. We have to select such a section of the population as can represent the whole population in all respects. We have to apply the statistics obtained from the sample to the whole population. The application of sampling statistics depends upon certain conditions of sampling. If these are not satisfied, standard errors, no matter how accurately computed, may give wrong impressions. At best, they give us only estimates from which we can make decisions and draw conclusions, never with complete conviction but with various degrees of assurance. This raises the question, as to what process of sampling should be followed to obtain the desired sample.

METHODS OF SAMPLING

Various techniques have been devised for obtaining a sample which will be representative of its population. Unbiased sample is the most adequate one. The adequacy of a sample will depend upon our knowledge of the population as well as upon the method used in drawing the sample. Commonly used sampling methods are as follows:

(1) Random sampling.
(2) Stratified-random sampling.
(3) Incidental sampling.
(4) Purposive sampling.

The discussion of each method in details can be found in any good book on statistics. Here is given a very brief description of each method.

(1) Random sampling: The term 'random' is often misunderstood. It does not imply that the sample has been chosen in an offhand, careless or haphazard fashion. Instead it means that we rely upon a certain method of selection called "random" to provide an unbiased cross section of the larger population. The criteria for randomness in a sample are met when,

(a) every individual in the population has the same chance of being chosen for the sample, and
(b) when the selection of one individual in no way influences the choice of another.

(2) Stratified-random sampling: This procedure should be applied when the population is composed of sub-groups of different sizes. The total sample should include individuals drawn from each stratum in accordance with the sizes of the sub-groups. Within each stratum the sampling is random.

Stratified sampling is illustrated in the standardisation of the 1937 Stanford-Binet Scale in the course of which about 3,000 children were tested.

Stratification is a common procedure introduced in sampling to help to prevent biases and also to assure a more representative sample.

(3) Incidental sampling: The term incidental sample is applied to those samples that are taken because they are the most available ones. Such a sample is often called "accidental".

Many a study has been made in psychology with students in classes of beginning psychology as the samples merely because they are most convenient. Such casual groups rarely constitute random samples of any definable population. And generalisations based upon such data are often misleading.

(4) Purposive sampling: A purposive sample is one arbitrarily selected because there is good evidence that it is very representative of the total population.

SIZE OF THE SAMPLE

The criteria of a good sample are: (1) It should be as representative of the total population as possible. (2) It should be as large as possible.

If we are compelled by the force of circumstances to select between the two, our choice must fall on the first one without any hesitation.

Suppose a population is composed of ten different strata. A sample of 500 drawn from the population including 50 individuals from each stratum in the population, is more acceptable than a very large sample of say 2000 individuals drawn from only five different strata of the whole population. This is because in the first case, though a sample is small comparatively, it is more representative, while in the second case, the sample is no doubt much larger but it is not at all representative of the population as a whole.

The dependability of M or σ is contingent upon the size of the sample upon which the SE is based. SE's vary inversely as the square root of sample size so that the larger the N in general the smaller the SE. Considering this fact, the sample should be as large as possible provided the representativeness of the sample is not affected negatively and it is also possible administratively and economically.

THE SAMPLE FOR THE PRESENT WORK

As is discussed earlier, this test is designed to administer to the prospective teachers to measure their aptitude for teaching. Arts and Science graduates are most generally appointed as secondary school teachers. Thus they form a population of prospective school teachers. A sample for standardising this test should, therefore, be drawn from this population.

As it was shown while discussing pilot testing in the previous chapter, the trainees in the various training colleges in the states of Maharashtra and Gujarat (old Bombay State) would form a very good sample for our purpose.

An analysis of the trainees in any college will show that they come from almost all parts of the state. They come from rural as well as urban areas of the state. The trainees belong to different religions, communities and castes. They represent both the sexes. Economically, also, they represent about 95 per cent of the population. Only a few who belong to the rich class in the society come to this profession.

It is seen that trainees, from all different strata, as shown above, of our population, seek admission to the different training colleges.

As there are about eleven training colleges in Maharashtra and Gujarat, a fairly large sample also was possible

to be obtained.

Considering all these favourable factors in obtaining representative and large sample, the investigator decided to draw the sample for the present work from the training colleges in both the states. It will be seen that this enabled him to get a fairly large and stratified-random sample. We can, without the least fear of committing an error, say that this sample is a 'purposive sample' also. This sample can be called 'incidental' also, as it is most easily and readily obtainable.

Whatever name ultimately we may give to our sample, it is fairly certain that it is a representative one of our population.

The following is a list of training colleges in Maharashtra and Gujarat.

Maharashtra:

(1) Bombay: Secondary Teachers' Training College.
(2) Bombay: St.Xavier's Institute of Education.
(3) Poona : Tilak College of Education.
(4) Kolhapur: Shrimati Maharani Tarabai Training College.
(5) Sholapur: Dayanand College of Education.
(6) Satara: Azad College of Education.
(7) Nagpur: University Training College.

(8) Aurangabad: Government B.Ed.College.

(9) Akola: Government B.T.College.

Gujarat:

(1) Ahmedabad: A.G.Teachers' College.

(2) Baroda: Faculty of Education and Psychology.

(3) Porbandar: R.G.Teachers' College.

(4) Anand: Secondary Teachers' Training College.

(This was started from June, 1960, i.e. about a year after the testing was over).

According to his original plans, the investigator was to draw the sample from all the training colleges listed above. But he had to change his plans and to drop the plan of administering the test to the trainees in the colleges situated in the interior parts of Maharashtra for the following reason.

The testing programme was to be completed latest by the end of June, 1959, as this test was to be administered to the newly admitted trainees before they got any substantial knowledge regarding theory or practice in the college. The investigator personally administered the test at all centres. It was physically impossible for him to move from one centre to the other within this short span of time.

Anyway, the investigator feels that the sample which he has been able to collect is very fairly representative of the population and of statistically acceptable size also.

The table below gives a complete picture of the

composition of the sample obtained for the test administration for its standardisation.

Sr. No.	Name of the training College	Size of the sample in the beginning	Cases rejected because of incomplete answering	Cases rejected because of not appearing in the University examination	Final size of the sample
1	2	3	4	5	6
1	Faculty of Education and Psychology, Baroda	113	13	-	100
2	A.G.Teachers' College, Ahmedabad.	94	17	3	74
3	R.G.Teachers' College, Porbandar.	71	4	1	66
4	Secondary Teachers' Training College, Bombay.	101	6	1	94
5	St.Xavier's Institute of Education, Bombay.	82	0	4	78
6	Tilak College of Education, Poona	121	1	2	118
	Total of Each Column:	582	41	11	530

Column 3 in the above table shows that the test was administered to 582 trainees in all. The test was administered with directions to take all the time needed and to mark every item even if this meant guessing on some items. The

testees worked on the items in good spirit and apparently co-operated well, but the answerbooks were not inspected before acceptance and it was later found while scoring that 541 testees had actually marked an answer to every item. Thus as many as 41 answerbooks were rejected because they were incomplete. This is shown in column 4 in the above table.

Column 5 in the same table shows that eleven individuals had to be rejected as their university marks were not available. Eight of them did not appear in both the parts of the examination while of the remaining, two did not appear in part I examination and one did not appear in part II examination.

Thus, as shown in the last (6) column of the same table, the final sample consisted of 530 trainees.

The places where the training colleges are situated in the states of Maharashtra and Gujarat in India, are shown in the map of India on page No. 228 .

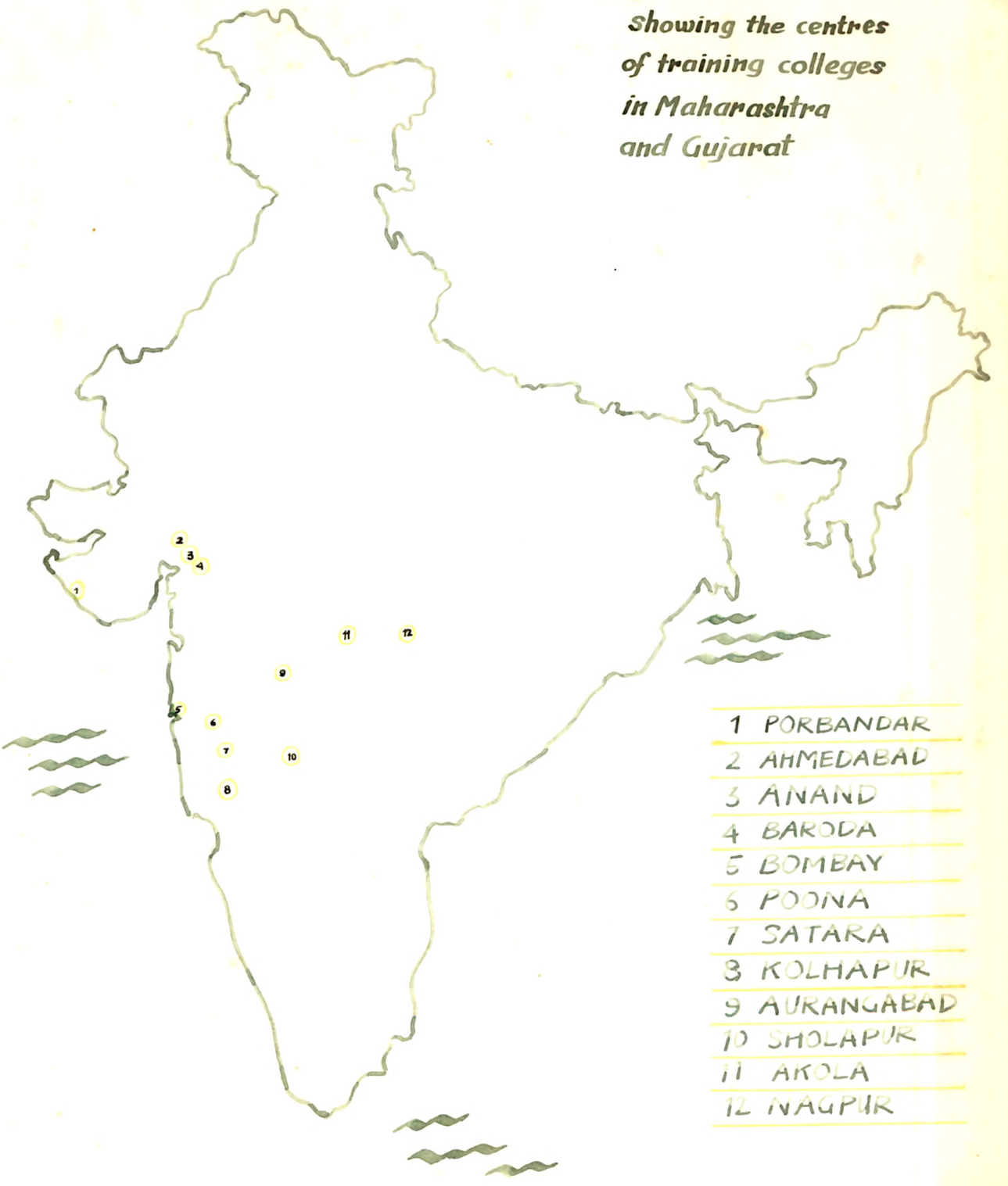
TEST ADMINISTRATION

Uniformity of procedure in test administration should be achieved so that the final evaluation of a given testee will be the same no matter where, when, or by whom he was tested. The point should be stressed that tests are standardised on the basis of a particular set of directions for administering. Comparisons with norms are valid only when exactly the same procedure is used in administering the tests that was employed

A map of

INDIA

showing the centres of training colleges in Maharashtra and Gujarat



1 PORBANDAR
2 AHMEDABAD
3 ANAND
4 BARODA
5 BOMBAY
6 POONA
7 SATARA
8 KOLHAPUR
9 AURANGABAD
10 SHOLAPUR
11 AKOLA
12 NAGPUR

when the norms were established.

Before we proceed to discuss the "standardised procedure" for administering this test, it is necessary to note some of the important "administration errors". Arthur E. Traxler[1] enumerates them as under:

(1) Incorrect timing of tests that involve a time limit.

(2) Lack of clarity in the directions to the testees.

(3) Failure to make clear to testees what they are expected to do about guessing.

(4) Variation in the physical conditions under which the test is administered.

(5) Too little or too much stress on motivation.

(6) Failure to control opportunities for chance or purposeful copying.

As regards the first three types of errors, the test-constructor should take utmost care, while constructing the test, to see that no chance is left for such types of errors to be committed. He should prepare clear directions regarding these points. The problems in the last three areas are largely

---

1 Lindquist, E.F., (Editor), "Educational Measurement", American Council on Education, Washington, D. C., 1955, pp. 330-332.

those of the test-administrator, although the test-constructor can aid the test-administrator materially by means of appropriate suggestions in the test manual.

As is already mentioned earlier, the investigator administered the test at all testing centres under his direct supervision. He tried fairly successfully to control the last three types of factors so as to reduce the errors, arising out of them, to the minimum.

THE FIELD WORK

Though it was planned formerly to complete the testing before the end of June, 1959, it had to be delayed by about three weeks as the investigator could not get required facilities to administer the test so early in June, when at some centres, even new admissions were being given and regular classes were not meeting. There were other administrative difficulties also.

The testing was, therefore, started in the beginning of the third week of July, 1959 and completed by the end of the last week of July, 1959. The investigator went personally to all the centres to administer the test. He took about eighteen days to complete the testing programme.

THE PROCEDURE FOLLOWED FOR ADMINISTERING THE TEST

(1) In the beginning, when all were seated, the answerbooks were distributed among the testees. Then, they

were instructed to fill in their names and the other items of information called for on the cover page of the answerbook. For this about two minutes were given.

(2) Then, the test-booklets were distributed among them with the instruction not to open it before they were advised to do so. When all the testees got the test-booklet, they were asked to read the directions on the cover page very carefully. For this, about five minutes were given, When each testee had done that the test-administrator gave some previously prepared additional instructions orally. He, then, asked if there were any questions from the testees regarding taking the test.

(3) He answered all the legitimate questions and asked the testees to open page 3 of the test-booklet and start answering items in sub-test I, with an instruction to read the directions at the top of the page and study the illustration given there carefully.

(4) They were not to ask any questions while answering the items. They were allowed to ask only the meanings of difficult words.

(5) The testees were to return the answerbook alongwith the test-booklet immediately after they had completely answered all the items in all the sub-tests. Before accepting the answerbook and test-booklet, he made sure that all the information necessary for identification and classification had

been entered in the blanks provided for the same on the front page of the answerbook and also that all the items in all the sub-tests were answered completely.

DIRECTIONS

The same directions that were fixed up at the time of pilot testing were retained as they were found quite satisfactory.

The basic instructions were printed on the front cover page of the test-booklet. The basic instructions needed to be supplemented by oral instructions from the test-administrator. These oral instructions were uniform for each group of subjects and were followed verbatim a typed manuscript.

The basic and the additional oral instructions are given on the page No. 170 in the previous chapter as well as in Appendix, E, and Appendix, I, respectively.

TIME LIMIT

No time limit was set for taking this test. The testees were instructed unequivocally in the beginning that they would be given as much time as they needed for taking the whole test. They were asked not to leave a single item unattempted even if that meant a little guessing in answering the item.

Thus, this being purely a power test there was no time limit at all. Even then, the investigator noted at all centres the maximum time taken by the slowest testees. The average of

these maximum times at all centres came to about 93 minutes. Thus it can be said that this test is approximately of 100 minutes. The test-administrator should plan his testing programme accordingly.

PHYSICAL SET-UP

To avoid the vitiating effects of fatigue on the testees' performance, it is advisable to give special attention to obtaining the best possible space, proper seating arrangement, adequate light, comfortable temperature and ventilation, and freedom from distractions. After all, these are all ideal conditions. An effort should be made to fulfil as many of them as possible. The investigator should honestly confess that he could not fulfil all the ideal conditions equally for test-administration at all testing centres.

At all centres, the testing was done during the first two periods of the college time. This facility enabled the investigator to administer the test when the testees were perfectly fresh.

MENTAL SET-UP

For all the successful psychological testing favourable 'Mental set' of the testees is a pre-requisite.

Some testees approach a test with complete lack of interest and considerable boredom and will not put forth their best efforts unless some outside motivating influence is -

present. Others become nervous and over-excited in a test situation and thus may find their performance blocked by their emotional state. The most desirable motivating conditions are, therefore, those which enable the largest number of individuals to turn in the best performance without undue emotional stress.

Arthur E. Traxler[1] feels,

> The personality and attitude of the examiner, of course, have a great deal to do with the motivation of the subjects.

The novelty in taking this type of an aptitude test played a great role of motivating the testees.

CORRECT ANSWERS

No change in correct answers fixed up at the time of scoring the pilot test was made. In Appendix, G, at the end of this treatise, a list of correct answers is given.

SCORING

The same scoring method, as was followed in scoring the pilot test, was followed here also. No change was necessary in the scoring pattern. The scoring pattern, thus, is a standardised one, for scoring, the present test.

No help was taken from any person for scoring. So there was no question of any scoring errors being committed.

---

1 Lindquist, E. F., Op.Cit., p.345.

The investigator personally assessed all the 530 answerbooks.

The scoring method has been discussed at length in the previous chapter. A brief summary of the same is given here just to maintain the continuity of discussion.

(1) Items in sub-test I are to be scored on a five point scale. The best answer is to be given 5 credits, while the worst is to be given 1 credit. All the credits are to be added together and the total is to be divided by five to get the final total out of 35. Each item in this test also ultimately gets a maximum credit of one. Thus the maximum obtainable score for this sub-test is 35.

(2) All the items in sub-tests, II, IV and V are to be marked as right or wrong, the right one getting one credit while the wrong one getting '0' credit.

(3) In sub-test III, the first five items are to be scored 1 or 0 as the case may be. Whereas for scoring items 6, 7, 8 and 9 a slightly different procedure is followed. In all these items four preferences are given. The best preference selected is to be given 4 credits. The next one is to be given 3 credits, and so on.

The scoring was done without the use of any formula

for correction for chance. The reasons for not using the correction formula for chance are discussed in the previous chapter.

KEY FOR MANUAL SCORING

No key for manual scoring of sub-tests I, II or III is possible, as the responses to be given to items in sub-test II are in the form of words or figures and credits are to be given in proportion to the merit of the responses to all items in sub-test I and items 6, 7, 8 and 9 in sub-test III.

For scoring sub-tests IV and V, a plain punched key is possible. The testees are to put a cross (X) on the corresponding letter of the alternative they have selected. The letters representing the distractors for each item can be printed on the sheets of light cardboard in exactly the same way as is done on the fourth page of the answerbook. Then the holes may be cut, in the places of letters showing correct responses, with the help of a punch. Then, crosses put on letters representing correct answers, can be counted easily through the holes.

A model scoring key for sub-tests IV and V is shown in Appendix, H, at the end of this treatise.

PREPARING DATA FOR STATISTICAL ANALYSIS

After the scoring of all the 530 answerbooks was over, the scores were transferred in a register specially prepared

for the same. Sub-test-wise scores were entered in the register. Then the total score of each testee was found out and ranks were assigned to the testees for the sake of identification of the answerbooks.

As the Means, SD's etc. were to be calculated for total-test score and for each sub-test score also, the relevant data for the same were prepared. These were to be found out for the total sample as well as for the unit samples also. Data for the same purpose were prepared. Thus all data for various types of statistical analyses were ready for use.

A LIST OF DIFFICULT WORDS

While constructing the test utmost care was taken to see that as simple a language as possible was used, lest it should be a 'language ability' test. Inspite of this, a little language intricacy has remained. The test-administrator should, therefore, explain some difficult points, if and when requested by the testees.

During the testing at different centres, the investigator was required to give the meanings of some difficult words. A list of these difficult words is given below. The meanings of these words should be given alongwith the instructions while administering the test.:

(1) Truants.

(2) Considerate.

(3) Code .

(4) Ethics.

(5) To connive at.

(6) To vindicate.

(7) Motivation.

(8) Boredom.

(9) Fatigue.

Wherever the investigator went to administer the test, he got full facilities and very valuable co-operation both from the college principal and staff members. He got much active and badly needed help from the college teachers in administering the test. The investigator takes this opportunity to express his deep sense of gratitude to all those who gave their very valuable and sincere co-operation in his work.

How can the investigator forget the testees at different centres? But for their very willing co-operation and active interest in taking the test, perhaps the investigator -'s work would not have become so successful and smooth. The investigator thanks them all for the same.