

CHAPTER IV

*
*
* * * THE PILOT STUDY AND THE
FINAL FORM OF THE TEST

The test, in its pilot form, with items framed on the basis of opinions of eminent creative persons as well as earlier literature, as discussed in the previous chapter, constitutes of five sub-tests. These are :-

- (1) Sentence Completion Test,
- (2) Uses Test,
- (3) Creative Writing Test,
- (4) Consequences Test, and
- (5) Problem Solving Test.

The first four sub-tests contain 10 items each, while the fifth test contains 9 items.

FIRST TRY-OUT

The first form of the test being ready, it became necessary to find out how much time it would take to administer it. Time is an important factor in administering

a test, for no test can yield truthful results if it involves fatigue. One of the most practical ways of finding as suitable time as possible was to administer the test to a small group of students and find how much time they would take to respond to each test. The investigator started with ~~to~~ giving 5 minutes to each item of the test Nos. 1, 2, 4 and 5, viz. Sentence Completion Test, Uses Test, Consequences Test, and Problem Solving Test and 10 minutes for each item of test No.3 viz. Creative Writing Test. Further, two minutes more for each item were given for reviewing the written paragraph in this Creative Writing Test. The distribution of time thus amounted to 5 hours and 15 minutes.

TIME ESTIMATION

50 minutes for Sentence Completion Test

50 minutes for Uses Test

100 minutes for Creative Writing Test

20 minutes for Reviewing the Paragraph

50 minutes for Consequences Test

45 minutes for Problem Solving Test

315 minutes.

Total Time: 5 hours 15 minutes.

This was obviously too long a period for administering a test. The investigator had, however, observed that the time allotted to each test was more than was necessary. She, therefore made another plan for the distribution of time, reducing it to $3\frac{1}{2}$ hours. Still, the experience again showed that this was also too long a period and it posed difficulties with regard to: (1) the element of fatigue that a $3\frac{1}{2}$ hour test would involve and the nature of the responses if the students who were made to sit or respond under condition of fatigue for such a long time; (2) the time that the schools could make available to students for sitting through the test; (3) the difficulty that may arise because of the inability of some students to appear for all the tests for various reasons, e.g. being on leave, etc. However, it was not possible to reduce the time further as that would affect the performance of the students and therefore it became necessary to find out ~~and~~ an alternative procedure to overcome these difficulties. This was achieved as under :

(a) Each sub-test was treated as a separate test which can be administered at different times. The sub-tests were administered thus: Sub-tests Nos. 1 and 2 were administered to one batch at the same time. This took about 45 minutes. Sub-test No. 3 was administered to a different batch of students which took about one hour. Sub-tests 4

and 5 were given to yet another batch which took about 1 hour and 45 minutes. These three batches were rotated alternately so that each took all tests at the end of three sessions. This method of the administration of the test removed chances of the same subject sitting for responding to the test continuously at a time for $3\frac{1}{2}$ hours under conditions of fatigue and thereby he worked hopefully in the best of his spirits. All the sub-tests were given to all the students by rotation. The investigator assumes that the five sub-tests would be a valid instrument for measuring creativity. The test items were so framed and planned for scoring that tests Nos. 1, 2, 4 and 5 would give scores on fluency, flexibility and originality of the subject, and test No.3 would test organization, originality, imagination and richness of the respondent. These qualities have been accredited as being the qualities of a creative person, as discussed in chapter III. It is, however, not necessary to repeat what the investigator means by these terms. She does not feel any need to review or reinterpret the existing meaning of these terms as established by authors of tests on creativity and therefore, she adopted them as such.

PILOT STUDY

After this first try-out for timing the test, the investigator administered the test to the main sample of

350 subjects personally, after establishing rapport which is necessary in a test of creativity. The students must be placed under conditions which are not likely to produce fear, tension or develop any kind of inhibition. They should on the contrary feel relaxed and easy so that their creative impulse is released and they are able to respond as creatively as possible. To achieve this, the following steps were taken :

(a) The investigator requested the teacher - the authority figure of the class - to leave the investigator alone after he/she had introduced her to the class.

(b) The students were given full assurance that the test had nothing to do with their examination results or promotions.

(c) An introductory talk was given to the class indicating the purpose of the test before the test was administered in a very informal and sympathetic atmosphere.

(d) The accommodation for the administration of the test was paid special attention. Adequate lighting arrangements were made. Arrangements were also made to avoid distraction of attention wherever it seemed to exist. The test was held in class-rooms with which the subjects were familiar.

(e) The test was held in the same manner in all the schools. All necessary and possible precautions were taken to ensure normal conditions for each test.

(f) The whole test, as stated earlier, consisted of 5 sub-tests and involved answers of various types. Some items needed short spaces and others quite long. In the interest of efficiency and economy it was considered necessary not to adopt the usual practice of providing spaces for answers below each item of the test, but to provide the test sheet and the answer sheets separately to the subjects. The nature of the test and of responses did not require fixation of spaces. There was inherent in it the possibility that students would ask for additional paper. The tagging up of additional paper on to the responses of particular test would create much confusion and involve much labour at the time of scoring. The procedure of separating the test sheet from the answer sheets removed these difficulties and saved time and labour. At the same time it did not affect the students performance, adversely. It rather provided them with the freedom to answer each test in as detailed or synoptic a way as they liked. An additional advantage of this procedure was that it economised expenditure on paper. Otherwise, optimum number of pages would

have to be provided which would considerably add up to the cost.

(g) The handy and easily manageable sheets were also intended to remove from the mind of students the apprehension of being subjected to a long and tiresome test. Also, the optimum spaces given might have suggested that the response is to suit the length of the space provided. Such a psychological position would have come in the way of a free acceptance by the students of responding freely to the test item.

(h) To facilitate full understanding, detailed instructions were given on the cover of the test sheets. Instructions were also given for each test stating the way in which answers were to be recorded. The instructions were made as easy to understand as possible. Besides written and oral instructions, those students who faced any special difficulty were offered necessary explanations so that they could understand the tests fully.

(j) The investigator tried to restore confidence to students in their ability to respond well. They were told that they are unique special individuals and their uniqueness should be reflected in their answers. This served also as a motivation.

DESCRIPTION OF THE SUB-TESTS

Below is given a very synoptic description of each of sub-tests.¹

Test No. I: Sentence Completion Test

This test constitutes of ten incomplete sentences, five of which are two-word and five three-word sentences. The subjects are expected to fill in the blanks in as many different ways as they can. Following instructions are given:

इस परीक्षण में कुछ अपूर्ण वाक्यों को पूर्ण करना है। इनकी पूर्ति विभिन्न प्रकार से की जा सकती है। इन वाक्यों की पूर्ति आप जितने भी भिन्न प्रकार से कर सकते हैं कीजिए। परन्तु प्रयत्न कीजिए कि प्रत्येक पूर्ति से बनने वाले वाक्यों का अर्थ अलग-अलग हो। वाक्य पूर्ति के लिए एक से अधिक शब्दों का प्रयोग किया जा सकता है। वाक्य किस प्रकार पूर्ण करने हैं दिये गये उदाहरण से स्पष्ट हो जायेगा।

The test is designed to measure fluency.

1 Test booklet is given in Appendix C.

Test No. II: Uses Test

This test constitutes of ten ordinary objects used in day-to-day life. The subjects are expected to give as many uses of each object as they can. The instructions given for the test are as follow:

नीचे कुछ साधारण सी वस्तुएँ दी गई हैं। ये कई तरह से प्रयोग में लाई जा सकती हैं। इनमें हर वस्तु को कितने भिन्न प्रकार से प्रयोग में लाया जा सकता है लिखिये। नीचे दिये गये उदाहरण से स्पष्ट हो जाएगा आपको क्या करना है।

The test is scored for fluency, flexibility and originality.

Test No. III: Creative Writing

This test constitutes of ten topics which have a potential for creative expression. The subjects are expected to write a paragraph on each. The length of the paragraph is not specified. The instructions given are as follow:

नीचे कुछ विषय दिये हैं। आपको इनमें से प्रत्येक विषय पर कुछ पंक्तियाँ या एक अनुच्छेद लिखना है। अनुच्छेद () में ऐसे विचार प्रकट करने का प्रयत्न करें जिन्हें कक्षा में अन्य विद्यार्थी न कर सके। प्रयत्न करें कि आपके उत्तर में कल्पना की उड़ान हो।

The test is scored for organization, originality, imagination and richness.

In the second part of this test the subject is expected to review his own paragraph. The instructions given are:

आप, अपने ही लिखे इन परीक्षकों, को फिर पढ़ लीजिए और क्रमानुसार लिखिये कि इनमें क्या कमियाँ रह गई हैं। अथवा इनमें क्या क्या और लिखा जा सकता था।

Test No. IV: Consequences Test

The test constitutes of ten improbable situations. The subjects are expected to give as many consequences as possible if the situation could become probable. Following are the instructions:

नीचे कुछ असामान्य दशाएँ दी हुई हैं। आप कल्पना कीजिये कि यह घटित हो सकती है। आपको इनके परिणाम, जितने भी अधिक से अधिक आप सोच सकते हैं लिखिए। निम्न उदाहरण से स्पष्ट हो जायगा कि आपको क्या करना है।

The test is scored for fluency, flexibility and originality.

Test No. V: Problem Solving

This test constitutes of nine problems. The subjects are expected to give as many solutions as they can to each problem. The instructions given are as follow:

नीचे कुछ समस्यात्मक स्थितियाँ दी गई हैं। आपको इनका हल सोचकर लिखना है।
हर समस्या के बहुत से हल हैं। इसलिए आप जितने अधिक से अधिक हल
सोच सकते हैं उन सबको पृथक् पृथक् पंक्ति में लिखिये।

The test is scored for fluency, flexibility
and originality.

SAMPLE FOR THE PILOT STUDY

The test was administered to Delhi High School pupils. As stated earlier, Delhi has various categories of schools. There are (a) Government Schools, (b) private schools, (c) municipal schools, (d) public schools and (e) Denominational schools. The test was administered to subjects studying in different categories of schools. Five schools were chosen for the pilot study. These included :

- (a) Two boys' schools,
- (b) Two girls' schools, and
- (c) One co-educational,

Gross-wise the distribution was :

- (a) Two government schools,
- (b) Two private schools, and
- (c) One municipal school.

Selection of pupils was done according to their age. This representative group was selected from the age range of 14 years to 16 years. The number of subjects from different schools varied according to the population of the school, totalling 350 in all.

The names of the school and the number of pupils selected from each school are given in table No.2 .

Table 2: Showing the Number of Subjects

Sr.No.	School	IX	X	Total
1	Raisena Bengali Boys School	53	28	81
2	Govt. Boys' Higher Secondary School	37	25	62
3	Municipal Girls' Higher Secondary School	28	34	62
4	Govt. Girls' Higher Secondary School	37	25	62
5	Sardar Patel Higher Secondary School	43	40	83
Total:		198	152	350

SCORING PROCEDURE

The test was scored to measure fluency, flexibility, originality, organization, imagination and richness. The investigator has accepted these terms as these have been defined by experts in the field as well as her predecessors working in similar problems. As stated earlier she has not felt any need to review or reinterpret these terms. For the sake of convenience and ready reference, however, the measuring of the terms as understood by the investigator are noted as under.

Fluency is the number of relevant responses given to an item.

Flexibility is the number of shifts from one idea to another.

Originality is the statistical uniqueness of the responses.

Organization, imagination and richness have been made clear by Yamamoto with the help of his five point scale for each aspect. The details are given in the following scoring method. (The complete scoring system is given in Appendix D).

Fluency

Fluency being the total number of relevant responses, there was no need to develop any technique for scoring it.

Flexibility

Having administered the test to the above mentioned 350 high school pupils, their responses were noted. Scoring involved slight differentiation in the treatment of different tests e.g. one method had to be employed for test Nos. 1, 2, 4 & 5, and a different one for test No.3 .

Test number 1, 2, 4 & 5.- It was observed that responses to each item varied from 5 to 25. The total responses to 39 items of tests 1, 2, 4 & 5 as answered by 350 students were noted down. The responses, staggering in number were examined and it was found that a number of responses were similar and formed a category. All the responses were therefore examined from this point of view and most of the responses could be classified under one or the other category. Quite a large number of responses were such as could not be placed under categories. Since some could not be dubbed together to form other new categories, they were considered as independent responses. Thus, all the responses were placed under well defined categories or retained as independent ones.

In order to check whether the categories formed by the investigator were valid, it was necessary to get responses to each category judged by different judges. The judges were asked to see whether the responses put in a category were appropriate. They were given freedom, not only of pointing out the inappropriate responses in a category but also of forming new categories. The judges were persons of standing.¹

¹ List of judges is given in the Appendix E.

The investigator compared the categories formed by the judges with the ones she had prepared herself and it was found that the judgments offered almost the same categories and the same range of responses in each category. On the face value of the ranges thus acquired, the investigator did not find it necessary to find out inter-judgment reliability. The categories were therefore, accepted as such.

Flexibility score thus represents the number of shifts from one category to another, i.e. the number of different ideas in this case.

Originality

For scoring originality, the frequency of each response was found out. This again involved going through about 70,000 responses in all, and noting down the frequency of each response. The responses then, were given weightage according to their frequency.

While scoring the items it was found that sub-test No.1 i.e. Sentence Completion test scored very poor. The responses that were received did not give any indication of creativity. It was obviously not necessary to retain such a test mainly because it would not yield any fruit. It was, therefore, discarded. The investigator had to decide the

weightage that she might give to each response. She followed the procedure adopted by Dr. Torrance¹ in scoring originality for his test of creative thinking. According to him, "a score of zero is assigned to all responses given by 5% or more of the respondents; a weight of one is given for responses given by subjects from 2 percent to 4.99 percent of the respondents. Responses not included in the above criterion are given weight of "2" if they show creative strength (i.e. if they require intellectual energy beyond what has been learned, practical and habitual, and result in responses that are away from the obvious and common place). "¹

The weightage thus given by the present investigator was as under:

Two marks were given to responses with frequencies of 2% or less than 2%.

One mark was given to responses with frequency between 2% to 5%.

1 Torrance, E.P.: Torrance Tests of Creative Thinking Manual for Verbal Test Form B, Personnel Press Inc., New Jersey, 1966, P.49.

Zero mark was given to responses with frequency more than 5%.

All the responses upto 5% frequency i.e. responses having weightage of one and two marks were considered to be original. This list of original responses, derived as above, was given for further check of judgment to 15 judges. The judges were expected to tick mark the responses they considered original. The criterion for acceptance of a response was that 8 out of 15 judges should agree on a response as original. It is interesting to note that most of the responses derived from frequency basis were agreed upon. Some of the responses that were thus removed carried weightage of 'one' only.¹

Sub-test No.3 contains items which were scored for originality, organisation, imagination and richness.

Scoring was done on the basis of Yamamoto's scoring system.² The complete system is divided into five parts (aspects), viz. organization, originality, imagination, sensitivity and richness. These aspects are measured on a five point scale

1 The list of Original Responses is given in the Appendix F.

2 The Yamamoto Scoring System is given in the Appendix G.

and the response to each point is given 1 or 2 mark according to the scorability of the response. It may, however, be stated, that sensitivity as scored by this scoring system, was not relevant for the investigator's test. It was therefore, not included.

The scores thus obtained on the tests with respect to different aspects were then item-analysed.

ITEM ANALYSIS

The investigator had by now responses of 350 students duly categorized and with frequencies of each response found out. It was now necessary to find out which items would show the most creativeness in an individual. This was done by analyzing all the items. The analysis was done as follows.

Different types of scores from the five tests were summed up. For example, scores for fluency on all the 10 items of 'Uses Test' were noted down. These scores were added to get a total fluency score. Scores for flexibility on all the above mentioned 10 items of the 'Uses Test' were also noted down. These scores were added to get a total flexibility score. And in the same manner, scores for originality on all the above 10 items of the 'Uses Test' were noted down. These scores were added to get a total

originality score. This procedure was repeated for 'Consequences Tests' and 'Problem Solving Test'. Scores for organization, originality, imagination and richness in 'Creative Writing Test' were also computed in the same fashion. Thus, 13 tables for scores on fluency, flexibility, originality, organization, imagination and richness were computed. This involved scoring of total 39 items (tests No.2, 3, 4 & 5 with no. 1 dropped), each for 350 subjects.

It is perhaps necessary to explain why the investigator did not adopt the usual statistical procedure. The two most popular methods of analysing items are finding out difficulty index and discrimination index.¹ Difficulty index was not computed because it is not relevant in a creativity test, as there are no wrong answers. Therefore, only discrimination index was computed. The normal procedure for doing so is to add the scores on all the sub-tests to compute a total score, arrange their total score in an ascending or descending order, locate upper 27% and lower 27% and compare the two to see which items discriminate the

1 Davis, F.B.: Item Analysis Data. Their Computation, Interpretation and Use in Test Construction. Graduate School of Education, Harvard University, Cambridge, Massachusetts, 1946.

higher score from lower score. This method, as it is, could not be applied to the present test. The reason being that each item had unlimited number of responses and each response was being scored three times, once for fluency, second time for flexibility and third time for originality. There were no single scores to be added to make a final score. Thus, it became necessary to analyse first each item three times for fluency, flexibility and originality and see if these three aspects had any relationship with each other. For this reason, correlations between these three aspects were computed (using product moment method of correlation). On the whole, correlation between fluency and flexibility was found to be 0.78, between flexibility and originality 0.58, and between fluency and originality 0.39 . It was clear that fluency and flexibility were highly related, while fluency and originality were positively related but not highly. Flexibility and originality were again positively related but not as highly as fluency and flexibility. This relationship made it necessary to treat fluency, flexibility, originality as not one but three different aspects. Therefore, to find discrimination index an item had to be judged against the three criterion.

Thus, to find discrimination power of the items the correlation matrices for each test were computed for test Nos. 2, 4 and 5. This was done as follows :

Each of the tests, excepting test No.3 as stated above, have three aspects i.e., fluency, flexibility and originality. Each aspect has 10 items. Correlation was found between each item of an aspect say of fluency with its total of all the 10 items. This was repeated in the same Uses Test for flexibility and originality. This process was repeated for Consequences Test and Problem Solving test.

The Creative Writing Test (No.3) has been scored on the basis of Yamamoto's scoring system. The scoring was done in the same fashion as given above with the difference that each item was analysed four times (instead of three times) for organization, originality, imagination and richness. Discrimination index was found out by finding correlation of each item of an aspect, say organization, with its total of 10 items. This was repeated for the other three aspects, namely originality, imagination and richness. Thus, another four matrices were prepared.

In addition to this, inter-items correlations were also found out. This was done to see the relationship of items with each other. The objective was to select the item which had high correlation with the total of the same

type, but comparatively low correlation with other items.

Thus, thirteen ~~as~~ correlation matrices were prepared, three for Uses Test (one for fluency, one for flexibility and one for originality), four for Creative Writing Test (one for organization, one for originality, one for imagination and one for richness), three for Consequences Test (one for fluency, one for flexibility and one for originality) and three for Problem Solving Test (one for fluency, one for flexibility and one for originality).

RESULTS AND THE FINAL TEST FORM

Results show in general, that an item which has high correlation with the total for fluency also has a similar high correlation with the total for flexibility. For instance, item no.4 in Uses Test has the highest correlation with the total for fluency; the same item, i.e. item no.4 in Uses Test has also highest correlation for flexibility. This pattern is true for practically all the tests.

In comparison, items having high fluency or flexibility did not have as high a correlation for originality. For example, item no.4 in Uses Test did not have as high a correlation for originality as it had for fluency and flexibility. But an item which has high correlation with originality has similar correlation for

fluency and flexibility. This indicates that an item which is a good measure of originality is also a good measure of fluency and flexibility. For example, item no.8 in Uses Test which has the highest correlation for originality also has high correlation for fluency and flexibility. The tables, one for each test, given below present the detailed results of the correlations, arranged in order of their values on first aspect.

Table 3: Showing Correlations of Each Item

Sr.No.	Item No.	Fluency	Flexibility	Originality
1	4*	0.74	0.72	0.45
2	1	0.62	0.55	0.21
3	8*	0.56	0.55	0.56
4	6	0.52	0.50	0.41
5	5	0.45	0.40	0.23
6	2	0.40	0.38	0.11
7	3	0.38	0.37	0.31
8	9	0.28	0.10	0.13
9	10	0.26	0.08	0.02

The table shows that item no.4 (sr.no.1) has highest correlation for fluency (0.74) and flexibility (0.72), while

item No.8 (Sr.No.3) has highest correlation for originality. Consequently both the items i.e. item no.4 and item no.8 were selected for the final test.

Table 4: Showing Correlations of Creative Writing Test

Sr.No.	Item No.	Organization	Originality	Imagine	Richness
1	2*	0.71	0.56	0.53	0.71
2	3*	0.68	0.62	0.57	0.65
3	5	0.57	0.34	0.48	0.50
4	4	0.57	0.28	0.44	0.60
5	1	0.55	0.55	0.53	0.65
6	7	0.45	0.46	0.32	0.46
7	6	0.41	0.09	0.18	0.43
8	10	0.16	0.19	0.00	0.53
9	9	0.12	0.22	0.00	0.15
10	8	0.12	0.10	0.46	0.33

The table shows that item no.2 (Sr.No.1) has highest correlation for organization and richness (0.71), while item no.3 (Sr.No.2) has highest correlation for originality and imagination (0.62 and 0.57 respectively). Accordingly, these two items i.e. item no.2 and 3 were selected for the final test.

Table 5: Showing the Correlation of
Consequences Test

Sr.No.	Item No.	Fluency	Flexibility	Originality
1	6*	0.70	0.70	0.38
2	5	0.67	0.66	0.32
3	8	0.67	0.61	0.25
4	7	0.67	0.51	0.16
5	2	0.64	0.63	0.26
6	4	0.61	0.59	0.19
7	3	0.48	0.59	0.39
8	1*	0.48	0.47	0.50
9	9	0.44	0.32	0.07

The table shows that item no.6(Sr.No.1) has the highest correlation of 0.70 on fluency and 0.70 on flexibility and 0.38 on originality. Item no.1 (Sr.No.8) has the highest correlation of 0.50 for originality and 0.48 and 0.47 for fluency and flexibility, respectively. These two items i.e., item no.1 and item no.6 were, therefore, selected for the final test.

Table 6: Showing the Correlation of
Problem Solving Test

Sr.No.	Item No.	Fluency	Flexibility	Originality
1	5*	0.69	0.55	0.29
2	2	0.66	0.39	0.44
3	4	0.59	0.54	0.18
4	3	0.59	0.49	0.40
5	6	0.54	0.34	0.19
6	1*	0.50	0.45	0.51
7	8	0.49	0.17	0.01
8	9	0.47	0.26	0.01
9	7	0.42	0.31	0.16

The table shows that item no.5 (Sr.No.1) had the highest correlation of 0.69 for fluency and 0.55 for flexibility and 0.29 for originality. Item no.1 (Sr.No.6) has the highest correlation of 0.51 on originality and 0.50 and 0.45 for fluency and flexibility, respectively. Consequently, these two items i.e. item no.1 and item no.5 were selected for the final test. Thus, 8 items in all were selected.

The investigator began the construction of the test with 49 items, ten each for the 1st four tests and nine for the 5th test. Ten items of Test No.1 were discarded at the very outset due to very poor response. Item analysis showed that 8 out of remaining 39 items, statistically analysed, were the most potential. These were consequently retained and the rest discarded. With this procedure emerged the final test form¹ consisting of four sub-tests with two items in each.

¹ The final test form is given in the Appendix H.