That man should learn so much of the world around them without learning more about themselves is quite surprising. Psychological tests came into general use, two thousand years after the technique for measuring the circumference of the earth was developed.

It is probable that the complexity of the human individual delayed the development of psychological measurement and one need not seek another reason for the slow progress in the field. The philosopher Kant was of the opinion that it would not be possible to have a science of psychology because the basic data could not be observed and measured. Modern psychologists have disagreed with Kant and ventured to define psychology as a science of behaviour, which can be observed and measured quantitatively in an objective and scientific way. However, they would agree with Kant to the extent that only certain kinds of psychological phenomena are open to observation and measurement.

Intelligence is one of such phenomena. It is the field, most widely explored. But the 'yardsticks' are still very crude. It is true, that there are serious and continuous attempts to refine them but it is doubtful whether the refinement would any day reach perfection. Cronbach remarks that we have come a long way from Wissler's trying to correlate college success with speed of cancelling a's, and from Binet's wondering whether psychologists had anything to contribute to the identification of mentally inferior children (Cronbach, p. 189). But, looking to the complexity
of the human individual, it is not pessimism to believe that measurement in psychology can never hope to reach the precision and accuracy found in the physical sciences.

However, this expectation of a crude 'yardstick' is no excuse for the test maker to make a light job of his work. It is on the contrary, a signal to take all possible precautions so as to minimize the crudity of the instrument.

The present investigator took the signal and the first thing she did was to undertake an extensive study of related literature. The rich bibliography, appended to the present work is an evidence of this effort. This study enabled her to anticipate the difficulties and also to avoid them or to minimize them.

THE RANGE OF APPLICABILITY

It was decided to prepare a group test of intelligence to measure the intelligence of the pupils studying in standards V, VI and VII. But it was felt that a more precise description of the population to be tested was necessary before commencing to design the actual test items. The reason for this lies in the very nature of mental tests.

Mental tests are defined as instruments for the measurement of individual abilities or types of behaviour, with maximum emphasis on differences due to original nature rather than to training or environment (Freeman, F. W., p. 2). Mental capacities cannot be seen, felt, heard or measured in any direct fashion. What the psychologist measures is the manifestation of capacity in action or in behaviour. These actions
or behaviours are influenced by schooling and environment i.e. experience. The problem before the test maker is to find out how far the differences between individuals are due to differences in their native capacity rather than due to differences in the training that they have received. It is a difficult problem. The test maker must find activities in which previous experience is a negligible factor. But such activities are not manifestations of complex mental abilities. The test maker solves the problem by limiting the variety of individuals to be tested and by selecting activities which are largely common to their experience.

In fixing the range of applicability, the following three factors were considered.

(i) Schooling
(ii) Age
(iii) Cultural background

Schooling:

It was decided to make the test applicable to boys and girls studying in standards V, VI and VII of the secondary schools in Gujarat. It was not thought proper to include the pupils studying in standards V, VI and VII of primary schools, though the same syllabus was followed in primary schools also. There is a very significant difference in the general tone of the two types of schools. Most of the primary schools are highly influenced by the principles of basic education; many of them still resemble, somewhat, the indigenous schools of India. On the other hand, the secondary schools which have grown out of Macaulay's English Education, are under a very strong impact of Western Education. The difference in the quality
and type of teachers and hence in the methods of teaching is very very obvious. Besides, the school populations also differ considerably in socio-economic matters. Ambitious parents, who are serious about their children's education, send them to secondary schools as soon as they pass the fourth grade of the primary school. These are also the parents who can afford to pay fees. The education in the ordinary primary schools is more or less free. Hence, pupils in secondary schools generally come from a comparatively higher socio-economic level and are also better motivated for academic progress.

Age:

It is customary to fix an age range to which an intelligence test is to be made applicable. This is absolutely inevitable when the entire population of certain age groups is to be included and when the applicability is not limited to particular school grades. In the present case, the applicability of the test was limited to standards V, VI and VII of secondary schools and hence it was thought more advantageous to apply the test to all the pupils found in these standards. Separate norms for each age group were to be fixed and so the inclusion of all the age groups found in these standards was not at all likely to affect the norms in any way. Hence, the present investigator thought it proper to decide the problem experimentally, rather than limit the applicability of the test to particular age groups.

Cultural Background:

The last problem to consider was the differences in cultural background. None can afford to ignore them in the days when local norms are being highly talked of and strongly recommended. The places in Gujarat
can roughly be divided into three types of cultures - rural, semi-urban and urban. The present investigator does not deny the differences and is of the opinion that separate tests or at least separate norms for the three types of cultures would have been better measures of intelligence than the present one to be used for all the three. But constructing three different tests was beyond the scope of the present work. In that case, it would have been necessary to limit the applicability of the tests to a single culture. But that attitude did not seem to be proper in view of the practical exigencies of the day. Intelligence tests for the lower grades are badly needed in all the secondary schools. The way out was to design a test based on the experiences, common to all the three cultures. This task, though difficult, did not seem to be impossible. The three cultures are more similar than distinct. There is lot of overlapping. The rural and the urban areas differ mainly in their agricultural and industrial bias. The semi-urban culture is a curious mixture of the two. But, it is no shock to find cows and stray dogs walking down leisurely along the high ways of the capital of Gujarat and electricity and use of machines in the remote villages. Besides, the secondary schools in the rural and the semi-urban areas are not different from or inferior to schools in the urban areas, as a whole. It is erroneous to class all urban schools as superior. There are certain very good schools but that is not the whole story. In fact, quite a few schools in the cities are poorly housed, staffed and equipped and good many of them are in no way superior to their counterparts in villages and towns.

Having decided to use a common test for all the three cultures, the investigator thought of the possibility of separate norms. The first difficulty to be faced was a precise definition of each culture. A culture
is a complexity which does not yield to a universal definition. Any definition of culture, however precise and detailed is bound to be subject to various interpretations. In that case, the very basis of this objective measure would have been dependent upon the subjective interpretations of the users of the test. In this connection, Freeman writes:

It would be exceedingly difficult to find a method of grading social environments so as to apply norms to them. Furthermore, many gradations of environment could be found and the same individual is subject to the influence of more than one environment. For example, his home environment may be of one sort and his school environment another. These complications and difficulties seem to make it inadvisable to create norms for social groups.

(Freeman, F.N., p.320)

A further difficulty in the construction of local norms is pointed out by Cronbach. He remarks that local norms change from time to time owing to population migration and changing school policies (Cronbach, p.123.) The remark does not fully hold good for us. The school policies, if at all they change, will change everywhere. Such is the set up of our educational system. But the fact of population migration cannot be ignored. There is a marked tendency towards quick urbanization; communication facilities create possibilities for a greater and greater fusion of all the three cultures. And, frequent revision of norms is not a possibility.

The last difficulty realized, was about the size of the sample. If local norms were to be established, it would have been necessary to base the norms on a smaller number of pupils, for the investigator could not have afforded to test more than ten thousand pupils on the whole. The decision was taken on the basis of Freeman's advice:
The desirability of using local averages rather than general norms is particularly great in those cases in which the general norms have not been established on a large number of cases selected at random.

(Freeman, F.N., p.320)

Hence, it was decided to establish common norms for all the three cultures so that they could be based on a large population, representative of all the three cultures which would not change appreciably in the near future.

THE TYPES OF TESTS

In selecting the types of tests, the test maker has to consider the following three aspects:

(i) Testing in group or individually.
(ii) Sampling the behaviour to be tested.
(iii) Selecting the media for testing the sample behaviour.

The first aspect was already decided upon while undertaking the work; the remaining two are discussed in this section.

Sampling the Behaviour to be Tested:

A psychological test is defined as "..... essentially an objective and standardized measure of a sample of behaviour" (Anastasi, 2, p.22). In so far as such a sampling of individual behaviour is essential, psychological tests do not differ from those in any other science. In this respect the psychologist proceeds in much the same way as the chemist, the geologist or the agronomer. All these analyse a sample, or in fact, a sample of samples.
The mental ability to be measured is bound to manifest through various behaviours. The psychologist cannot and need not measure them all. While discussing the sampling of traits and functions, Freeman writes:

Any given test measures a limited aspect of the person being examined, though some tests are much more restricted in scope than others. It is essential, therefore, that the test builder define the aspect, or aspects, he proposes to measure. After doing this, he must develop a series of test items that will best sample the traits or functions with which his test is concerned.

In developing a psychological test, it is impossible, and in fact unnecessary, to use an unlimited number of items. It is not necessary to attempt to present the individual being tested (called the "subject" or the "testee") with problems that will ascertain his responses for every conceivable situation involving a given trait or function. It is sufficient to get an adequate sampling of responses in a particular area or range of behaviour, the assumption being that the sampling is representative of the whole.

(Freeman, F.S., p.4)

With regard to defining the aspect to be measured, the present test maker has accepted mainly, the Spearman approach to intelligence. This approach raises a problem. If a general factor is common to all mental activity, the test maker can, with advantage, select an activity, having high demands upon general intelligence and save himself the labour of having to construct various types of tests, as are usually found in a battery designed as a measure of general intelligence. The logic behind the principle is very sound. Attempts have been made to construct tests on this principle. Reasoning, analogy and classification are functions supposed to be highly saturated with g. We have Burt's Reasoning Test, Miller Analogies Test and the Non-language Multi-Mental Test by Terman and others. But such tests do not serve the general purpose. They work well when differentiation among individuals in a rather highly selected group is desired. Besides, so
far as the present test was concerned, it was feared that the same type of activity was likely to bore the pupils and diminish the influence of interest which was inevitable in motivating them to put in their best.

While justifying the use of a variety of tests, Freeman remarks:

The need for this variety is due, not so much to the fact that different mental processes are measured by them as that each of them measures the mental activity only as it appears in certain concrete operations of thought, and that these different concrete operations are conditioned partly by their material embodiment, as well as by their form.

(Freeman, F.N., p.254)

Practice has proved that a combined score of a series of tests is more reliable and a better predictor than the score of each subtest (Freeman, F.N., p.248 and Freeman, F.S., p.177). This is so, probably because daily or periodic fluctuations in performance are compensated for, through the testing of a number of representative functions (Freeman, F.S., p.177).

Hence, it can be concluded that an item pool composed of a variety of types of tests and material would be a better sample of the item universe to be tested.

Selecting the Media for Testing the Sample Behaviour:

Four main types of media have been made available by the veteran test makers of the West.

(i) Using language and numerical symbols.

(ii) Using pictures, diagrams and non-language symbols, with a view to avoiding the use of language.

(iii) Using manipulation of objects, with a view to avoiding both the language, as well as, the paper-and-pencil work.
(iv) Using motion picture film or television, with a view to achieving more accurate standardization.

To think of using the fourth medium in India, under the present circumstances, would mean planning an improbability, though not an impossibility. The third type, the performance type, has traditionally been restricted to individual testing. The last two media could, thus, be summarily dismissed. The present investigator had then, to choose between the first two media only. But it was a really tricky situation. The paragraphs that follow, give the reader an idea of the situation, and also of how the investigator attempted to meet it.

Right from the days, since Binet constructed his scale, verbal tests are being frowned upon by a large majority of test critics, test users and test makers. They are being severely criticised for the use of language which is mainly the product of schooling. And yet, it is the medium, most widely used. Practice has established its efficacy so much, that the psychologists tolerate it as a necessary evil.

Non-language tests too, have their place in psychological testing. The use of such tests is recommended for illiterate, foreign-born, bilingual deaf and otherwise linguistically handicapped subjects. It is also thought quite legitimate to use such tests, when the subjects are suspected of defective reading ability or inadequate schooling owing to frequent illness, absence or emotional disturbance. They are thus used to check the doubtful results of verbal tests for they correlate favourably with the verbal tests, the correlations ranging between the .50's and .60's. However, it is usually never thought advisable to use the non-verbal tests in preference to the verbal ones. Cronbach believes that most of the non-verbal items do
not call for very complex mental processes and that if "higher intelligence" involves ability to do abstract thinking, it is hard to conceive of an adequate high level test that does not involve vocabulary and concepts. He further states that their empirical validity is only moderate for they do not correlate highly with verbal tests or with school success and criteria of success outside school have not yet been developed (Cronbach, p.188). Freeman also supports the same view and adds that the validity coefficients fell in the .50's and .40's in the case of pupils in the later grades owing to the inability of the most available non-verbal tests to discriminate between individuals in the upper level of ability (Freeman, F.S., p.266).

Thus, language seems to be the principle source of the strength of the verbal tests but it is also its weakness. But as mentioned earlier this weakness is accepted as a necessary evil by the psychologists. Terman defends Binet's scale by replying to the critics that intelligence at the verbal and abstract level is the highest form, the sine qua non, of mental ability (Terman quoted by Freeman, F.S., p.127). It must be noticed that Terman defined intelligence as the ability to deal with abstract terms, and to do conceptual thinking. Another psychologist, Cronbach, while discussing whether Stanford-Binet scores are strongly weighted with verbal ability remarks that a child may do badly on the test because of poor schooling, but that this will also cause him to do badly in school in future (Cronbach, p. 114).

There should be no difficulty if intelligence tests are considered tests of scholastic aptitude. The stress in our culture on verbal education as the key to advancement has set the pattern for the tests. "Traditional 'Verbal' measures are still better predictors of the available criteria.
It may be that as performance in real life situations beyond school is studied more freely, the 'culture fair' type of test will come into its own." (Nunnally, p.225).

The question then is - is the present test meant to be a measure of scholastic aptitude? The answer has to be in the affirmative and in that case, the obvious corollary is that verbal tests would suit the purpose most. But the present investigator could not accept such a solution. She could not ignore Desai's experience (Desai, p.56). When he tried his battery with the pupils of standard VI, he found that most of them could not follow at all what was to be done in the tests of Proverbs, Series, Analogies and Story Completion. She feared that the situation was bound to be worse with pupils of standard V who had one less year of schooling i.e. only four years of schooling. Burt's warning too, could not be disregarded. In a foreward to Sleight Non-verbal Intelligence Test, he writes:

"Hitherto, most of these group tests have been couched in verbal form. They imply that the child is already able to read the instructions and to write his replies. With younger and duller children, however, reading and writing themselves involve special difficulties; and for this purpose psychologists have endeavoured to device what are called 'non-verbal' or 'performance' tests."

(Sleight - Manual of Directions, p.2)

Kelley also found that what one verbal group test measures is 92 per cent the same as what typical reading tests measure and is 90 per cent the same as what is measured by the typical informational school achievement tests. In addition to the language handicap, it was also felt that verbal tests alone might fail to keep up the interest of the pupils. Hence, for pupils with limited school experience, a group verbal test was sure to be far from
However, it was not decided to construct a test wholly non-verbal. The test maker was not prepared to let go completely the advantages of a verbal test. She, therefore, decided to include in the battery, such verbal tests as could be constructed within the limited linguistic attainment of the testees and such as could maintain their interest. The rest of the test was to be non-verbal. Thus, a mixed battery was planned. It was not a new venture. There seemed to be no risk. Wechsler had tried it in individual testing and made a name in the world. It was also tried successfully in group testing by Kuhlmann - Anderson, Sullivan and others.

All these have been successful for they have selected non-verbal tests which resemble verbal tests in the complexity of mental processes. Such tests are not impossible to find. In fact, on the basis of past experience, Wechsler thought that both verbal and non-verbal materials provide the most adequate and representative content, rather than either one alone. Hence, the Wechsler scales combine verbal and non-verbal materials within a single instrument in an effort to obtain the advantages, comparisons and contrasts yielded by both types of materials (Freeman, F.S., p. 157).

Some clarifications are necessary before closing the present section. The non-verbal tests to be used would differ from non-language tests which are administered with pantomime directions. It was not necessary; besides, verbal instructions work better than pantomime directions. Neither was the test to be culture fair. It is doubtful how far the tests that claim to be culture free, culture fair or cross cultural are really so. But it is not necessary to discuss the problem. It is enough to notice that
such tests use non-language symbols, abstract in nature. That sort of test material could not possibly be understood by the children so young and so inexperienced in taking tests. Besides, thinking of including such tests would have meant creating unnecessary difficulties. In so far as it was decided to include verbal tests, the cultural effect was included. It was thought that pictures depicting the life scenes within the experience of the normative population could be used with advantage as done by a number of test makers in the West. It was no use creating unnecessary complications by imitating Cattell's culture free test. These pupils could not be expected to think beyond their experience, through symbolic drawing which had no meaning for them.

**SELECTION OF THE TESTS FOR THE BATTERY**

The principles on which the tests for the battery were to be selected have been discussed elaborately in the foregoing paragraphs. They are enumerated below for clarity and convenience.

**Principles Followed in the Selection of the Tests:**

1. The tests should demand mental activities within the limits of the academic achievement and school training of the pupils of standards V, VI and VII studying in the secondary schools of Gujarat.

2. Those activities should be suitable to the intellectual maturity and interest of boys and girls between the age range 8.6 to 15.4 and above.
(3) The test material should be based on the common life experiences of the pupils, living in all the three types of cultures — rural, semi-urban and urban.

(4) The tests should sample adequately the behaviour to be tested.

(5) There should be variety in the type and material of the test.

(6) The battery should include both, the verbal as well as the non-verbal items.

(7) The verbal tests should not demand a high level of linguistic attainment and ability.

(8) The mental processes demanded by the non-verbal tests should as far as possible resemble those demanded by verbal tests, commonly used.

(9) The non-verbal tests should be safeguarded against over simplification and should be made capable of complex mental processes.

(10) Each test should conform to the Spearman approach to intelligence and hence saturated with g.

(11) The battery as a whole must be a good predictor of scholastic aptitude.

(12) The non-verbal tests need not use pantomime directions.

(13) The non-verbal items should not necessarily be 'culture-fair' but might be composed of pictorial material.

(14) The battery should be suitable for group administration.

(15) The method of answering should be simple and not complicated.

The task of selection was to be done in view of the above principles. Desai modelled his tests on the patterns of the Otis Advanced
Examination and the Army Alpha Tests (Desai, p.42).

The present investigator tried to look for a model but arrived at the conclusion that it would be more advantageous to select suitable types from various batteries rather than base them on one or two. Hence, she began the process of rejection, selection or adaptation. It is neither possible nor necessary to give an exhaustive list of the types that were rejected. But it was thought desirable to give reasons for having rejected some of the widely used types. Such a discussion can help to clarify the investigator's point of view and can also serve as an evidence for the efforts made by her.

Tests that were Rejected:

(1) Series: The series tests are found in both the media, numerical and non-verbal. The numerical form is very common. The test is based on perception of relations among a group of things. It is full of abstractness which is considered to be essential for measuring intelligence (Spearman and Jones, p.190). But the pupils of this age group do not seem to be capable of this degree of abstractness. Desai found that it did not work with the pupils of standard VI (Desai, p.56). The non-verbal form is found in Army Beta popularly known as the X-0 series. Probably because of the abstractness of the material, pupils do not follow what is exactly to be done. The population for the present battery cannot also be expected to have a sufficient mastery over numbers.

(2) Following Directions: The non-verbal form of this test is found in Mellen's Moray House Picture Intelligence Test. But the tasks are very simple and suitable to that scale, meant for the seven-year-olds. The verbal
form is used by Desai. He found the test highly (706) saturated with g
(Desai, p.194). However on studying this type carefully, it was felt that
such a test must necessarily be complicated. If the directions to be
followed are simplified there is nothing to be tested. In so far as the
directions are complicated it would be rather a test of reading ability
for the present group.

(3) Completion: The verbal variety demands of the testee to fill up a
gap or gaps in a sentence or a passage. The non-verbal variety is really
attractive. It is used by Wechsler and is also found in Army Beta, Kuhlmann
-Anderson tests and other Batteries. It appeared to be really a good and
suitable test but had to be rejected for it was feared that pupils unused
to tests would waste a lot of time in drawing the missing part of the
object.

(4) Rearrangement: This is another interesting variety. The verbal form
is known as jumbled sentences. The items look very cumbersome because of
the typical language structure of Gujarati. The case suffixes in Gujarati
hinder the construction of really sound items. Besides, the test is highly
verbal.

The non-verbal form requires the subject to rearrange parts of
a whole or to rearrange pictures in their correct order to present a
logical sequence of events. This test can work quite well when used in
individual testing as done by Wechsler. But the group form requires the
subject to indicate the rearrangement by numbering the parts or the
pictures. No simple method of answering could be devised and hence, the
type could not be included in the present battery.
(5) Matrices: This device has been exclusively used by Raven in his Progressive Matrices Tests which is a well known culture free test. It evaluates a person's ability to discern and utilize a logical relationship. Factorial analyses have shown the test to be largely a measure of g (with a small loading of spatial perception factor) which is described by the author as Spearman's education of relations and education of correlates.

The test was not included precisely because norms for Gujarati children on Raven's Progressive Matrices are being established by the B.M. Institute of Child Development, Ahmedabad. Besides, it was rather a very expensive affair to use the type. It is composed of series of test items of the same type and this does not lend variety.

(6) Reflections: This test is familiarly known as mirror images. Cattell uses it in the form of 'pool reflections'. The subject is required to identify the one of six drawings which represents the specimen drawing as it would appear in a pool image. This is a test of spatial perception based on a knowledge of scientific principle. Such a knowledge could not be taken for granted in the present group.

(7) Block Counting: This non-verbal type known as cube analysis in Army Beta, AGCT and Chicago Non-verbal Examination has been frequently imitated by many test makers. It is a space test and demands of the subject a three dimensional visualization. This type is used in India by the Allahabad Bureau of Psychology in their space relations test. As another better space test could be thought of, the use of this oft repeated test was avoided.

(8) Code Language: This type as found in English appeared to be really interesting and a very serious effort was made to construct one in Gujarati.
It was found that a satisfactory measure of the type could not be prepared. Vowel sounds are not indicated by separate letters in Gujarati as in English. The Gujarati alphabet does contain symbols for the vowel sounds but in writing they become a part of the consonant symbol. The test became very simple when the vowel and the consonant sounds were provided in combination and it turned out to be extremely complicated when separate codes for the two types of sounds were given. The investigator has arrived at this conclusion after having constructed the tests in both the ways shown above and tried on some pupils belonging to the standardization group.

(9) Family Tree: This type is often used for the present age group in the British tests. It arrested the attention of the present test maker. A test of the type was constructed but it failed completely in the first trial. The reasons of the failure were very obvious. The joint family system is fast breaking in India and hence, the children, especially those coming from the upper economic level are not familiar with the terms used to demote various relations.

(10) Memory: This type is very common in individual tests. Binet's tests include the testing of memory of several kinds. It is difficult to incorporate it satisfactorily in a group test. Such a test is found in Desai Group Test of Intelligence. The examinees listen to the story read by the examiner and then answer the questions based on the story. The ability tested is not mere retention but also attention or critical listening. (Desai, p.50). The pupils of the present age group could not be expected to be capable of the concentration and critical listening necessary in such a test.
Tests that were Selected:

Elimination and selection are two different terms but denote the same process. The process of elimination results into selection and vice versa. The selection that resulted included the following eight types.

1. Reasoning
2. Analogies
3. Story Completion
4. Matching
5. Absurdities
6. Similarities and Differences
7. Classification
8. Spatial Relations

Story completion is necessarily verbal and spatial relations non-verbal. The remaining six types can either be verbal or non-verbal. The descriptions of the tests, and the reasons for their suitability as well as for the use of the particular medium are given below. It must be noticed that the ten tests to be described are those used during the first empirical tryout of the battery (Appendix A and B). For convenience reference to the appendix is made before the discussion of each test. Their arrangement in the battery, titles etc. are given at the end of the discussion.

(1) Reasoning (Appendix A, Test 1): As mentioned earlier only verbal medium can be used for this type. It is directly based on 'Logical Inferences' found in Desai Group Test of Intelligence. The mental process required is reasoning or verbal analysis of a situation. There was a great temptation to include the test also because Desai found that it amused the pupils (Desai p.47). The only objection in view of the principles accepted in the construction of the present test was that this type did demand a considerable amount of reading ability. An attempt was hence made
to present the problems in as simple a way as possible with a view to minimizing the reading difficulty. However, the fact could not be denied and as will be mentioned later on, the test could not survive till the final tryout.

There are fifteen reasoning problems. It is a multiple choice type, for the testee has to choose the right answer from the four following each problem.

(2) Analogies (Appendix A, Test 3): There is hardly any battery of verbal group intelligence tests that does not include analogies since 1910, when Burt used it for the first time in intelligence tests. "It is the result of an attempt to humanize mathematics" (Ballard, 1, p.18). Desai remarks that it is the best example of the cognitive processes described by Spearman (Desai, p.48).

The type is less frequently found in the non-verbal form. Some of the pictorial analogies included in California Tests of Mental Maturity and Helaine's Moray House Picture Intelligence Test, look far-fetched. It was feared that it would not be possible to construct a sufficient number of non-verbal items which might neither be simple nor far-fetched. Besides, the investigator was deficient about being able to make the pupils follow, what was to be done.

It was decided to select the verbal type used by Desai. But in Desai Group Test, the position of the missing term is sometimes third and sometimes fourth. In the present test, it is always the fourth. This was necessary in view of the age of the group. Thirtyfive verbal
items of the type were coined and three were given as illustrations.

(3) **Story Completion** (Appendix A, Test 5): This verbal test is mainly based on Desai Group Test of Intelligence but the method of answering is simplified. In Desai's test the blanks and the groups of words to be used for filling in the blanks are numbered. Besides, some blanks (8) are not numbered and are not to be filled in. But the testee has to put the missing word mentally so as to build up the story. Such a process was bound to be complicated for this age group. Hence in this test no blanks were kept. Instead, three words are printed in bold type and the subject indicates his choice of the right word by underlining it. The test contains twenty five items and an additional one for illustration.

The test is partly a completion type. The only difference is that the subject has not to supply the answer from his own knowledge. The test, no doubt, is mainly a reading comprehension test but it also demands the use of logic. Pupils find the story interesting and the distractors amusing. This test has given the best results in the item analysis procedure and its g saturation in the whole battery is the highest.

(4) **Matching**: The word 'matching' popularly denotes the method of answering but it also suggests the mental process involved. Matching means associating rightly. To associate rightly the subject must detect similarities. Both the verbal and the non-verbal types are used in the present test as described below.

(a) **Verbal** (Appendix A, Test 4): This type was suggested by the Proverbs Test in Desai Group Intelligence Test. The pupils of standards V, VI and VII cannot be expected to know the proverbs so well. So this test requires them to match professions and things. The test demands
information but the information is such as they can easily obtain from their life experiences and is not dependent on schooling. The stress is on the right use of words.

Thirty items, excluding the two for illustrations, are given in three matching sets. In each set there are two extra names of things to prevent the last answer being obvious through elimination. For simplification, the use of numbers for indicating the answer has been avoided. Instead, the subject has to write the word itself.

(b) Non-verbal (Appendix B, Test 1) : There are twelve tables (including one for illustration), each having legs of different shapes. Below there are given fourteen numbered legs. The matching is to be indicated by writing the number of the leg on the table. The use of numbers is not avoided here as in the previous test. No other suitable method could be found. From experience it was also thought that the pupils would be able to manage the numerical symbols associated with pictures rather than with verbal symbols. The mental process required is, from perception and association.

(5) Absurdities : This type of tests amuse the pupils. It was first used by Binet in intelligence testing. "The detection of absurdities is one of the most ingenious and serviceable tests of the entire scale. It is little influenced by schooling, and it comes nearer than any other to being a test of that species of mother-wit which we call common sense" (Terman quoted by Ballard, 1, p.46). Both the verbal and the pictorial absurdities were tried in the preliminary tryout.

(a) Verbal (Appendix A, Test 2) : Various methods of using verbal
absurdities were considered. In Binet's absurdities the subject is required to say what is foolish or funny about the statement. If the answer is not satisfactory he is further asked why it is foolish or funny, Ballard (Ballard, 1, pp. 56-57), however, objects to Binet's telling the children that the statement is silly. He suggests that real absurdities should be mixed up with pseudo-absurdities so as to keep the testee in doubt. He should then be asked to write which statements are absurd and why they are so. In Chelsea Mental Tests Ballard has used absurdities in fool-proof form. He makes it a multiple-choice type test by giving below every absurdity four tries at saying what is foolish in it and requiring the testee to choose the best. The present investigator felt that the fool-proof form made the tests very easy. She preferred the absurdities being explained through reasons but she could not take the risk of making these pupils write a long answer for, they differed widely in their speed of writing. Such a procedure would also have made the scoring less objective. Hence she decided to mix up real absurdities with pseudo-absurdities and make it an alternate choice type test by requiring the testee to put a cross if the statement was absurd and put a tick mark if it was true. Fourteen absurdities along with thirteen pseudo-absurdities made twenty-seven items besides three for illustrations. The number of pseudo-absurdities is very large compared to that of Ballard who kept only four in a test of thirty items.

The present investigator is not an authority, competent enough to criticise Ballard, but it was her honest opinion that Ballard's procedure would misguide the testee. She also differed from Ballard in her method of scoring. Ballard did not score the pseudo-absurdities. But it was felt that the subject would be required to use his intelligence in deciding that
a particular statement which looked absurd was really not so. Hence, in
the present test all the items were scored.

But the test did not work. It was a miserable failure. The item
analysis results were erratic with regard to difficulty and discrimination.
Almost fifty per cent of the items had to be rejected and consequently
the test was dropped out of the battery. This failure is not difficult to
explain. A two option test usually does not give good results. Ballard
and Binet could be successful for they insisted on reasons being given.

(b) Non-verbal (Appendix B, Test 4) : The pictorial absurdities
are modelled after Revised Beta Examination and Mallow's Moray House
Picture Intelligence Test I. The pupils found them more amusing than the
verbal absurdities. Each of the thirtytwo items contains four pictures,
depicting objects or persons in largely similar situations. There is
something absurd in one of the pictures; the testee is required to detect
this and cross out the picture. Broad general similarity of situations
gives the clue to concentrate attention on a particular aspect and lend
to the detection of the absurdity. The mental process required is more or
less the same as the one in verbal absurdities, though there is one
significant difference. Verbal absurdities are almost always impossibilities
whereas the pictorial absurdities are impossibilities or often abnormalities
and hence improbabilities. The test was successful probably because it
was a multiple choice type and perhaps, also because, the pictorial
material interested the pupils and motivated them to think on the right
lines.

(6) Similarities and differences (Appendix B, Test 2) : This type of
verbal test often requires the subject to match words similar or opposite
in meaning. But it becomes purely a vocabulary test and hence the verbal form was not selected.

The non-verbal test is based on the one found in the Revised Beta Examination. It gives pairs of objects and numbers, some of which are similar and some different. Pairs of numbers were not used in the present test for it was likely to be a tedious task to the pupils. Forty-nine pairs of objects including the three used as illustrations were given. As the pupils worked on the test, it was felt that they were quite engrossed but this test was another huge failure during the item analysis procedure. The pictures in the present test are small but those in the Revised Beta Examination are smaller still. The only conjecture that the investigator is tempted to make is that the differences in the Revised Beta Examination are very obvious and it is mainly a speed test. In the present test the differences are very minute and purposely guarded from being obvious for the battery was being designed to be mainly a power test. The detection of such minute differences requires a very careful observation of which the pupils were not capable.

(7) Classification (Appendix B, Test 3) : This is another type which is a 'must' in intelligence test batteries. The verbal form was avoided for, the pictorial form was found to work well in many tests such as the Chicago Non-verbal Examination, Malone's Moray House Picture Intelligence Test I and Non-language Multi-mental Test by Terman and others. The authors of the last test, which employs only this single type throughout, state that their scale "..... is constructed to measure the ability to recognize and utilize relationships not among verbal symbols, but rather among pictorial symbols" (Quoted by Freeman, F.S., p.252). The present test contains
twentynine items in addition to two, used for illustration. Each item consists of drawings of five objects, four of which "belong" on the basis of some common relationship, while the fifth does not "belong". The testee has to identify and cross out the non-belonging item. In the construction of the test particular care was taken to see that this common relationship was not based on mere appearance but that it should be based on ideas. The testees have, hence, to pass through a complex mental process. The test is found to be highly saturated with 'g' and ranks second in this battery.

(8) Spatial Relations (Appendix B, Test 5): Such tests now-a-days are known as tests of spatial ability. They require the subject to reason about forms, or to recognize relations between them. These are either paper-and-pencil or apparatus tests. The question is whether it is legitimate to use such a test in a battery which is a measure of general intelligence. There is no general agreement but theory and practice both support such a use. In fact such tests were originally designed for the Army Beta test which measures intelligence non-verbally. They were also used by Stephenson (1931) in the battery whose intercorrelations, he attributed solely to g. (Stephenson quoted by Vernon, 2, p.66). Vernon, while summarizing the discussion on verbal and non-verbal factors in intelligence tests remarks, "Spatial or q tests are nowadays distinguished from intelligence tests, but there is no clear dividing line, and non-verbal g tests - abstract or pictorial - usually show a small spatial-perceptual component" (Vernon, 2, p.64). Space tests are largely used in intelligence testing and their g saturation is not less than that of any other non-verbal tests. The spatial factor is loaded with g like the verbal and the numerical factors.
These three factors, hence, create troubles for the constructors of differential aptitude batteries. During the factor analysis study of the General Aptitude Test Battery, developed by the United States Employment Service, it was found that a factor closely akin to Spearman's $g$ was common to the verbal, numerical and most of the spatial tests. (Anastasi, 2, p.380). In view of these facts, it was thought proper to include a space test in the present battery.

Different varieties of the spatial relations tests were studied. Those found in spatial aptitude tests and differential aptitude test batteries were of a complex nature; some of them needed three dimensional visualization. Such highly specialized tests cannot be expected to be good measures of $g$. But the types used in general intelligence tests looked to be quite suitable. In Army Beta the subject has to draw lines in square to show how the given pieces can be fitted to form the square. In the Chicago Non-verbal Examination also, the subject indicates how parts fit together to make a whole.

Some change in the above patterns was thought desirable. Most of them dealt with geometric figures. As mentioned earlier during the discussion of the media for testing, the present investigator is of the opinion that pupils of this age group cannot be expected to work well with material which has no meaning for them. The construction of the present test is influenced by her belief. Rather than talk in terms of triangles and rectangles, she preferred to use the word mirror. A part of the mirror is broken; this broken part is indicated by shading. The testee is required to identify the broken piece from among the four pieces given. There has, surely, been criticism from friends - mirrors do not break like
This! That is true but there is an answer—children can always distinguish between fact and fancy and enjoy indulging into fancies.

There are thirty three items including one for illustration. The mental activity demanded is visualizing and manipulating objects in space.

The preliminary design of the test was over. There were ten tests which contained 282 items. All of them could not be arranged in a single battery and administered on a single occasion. Hence the ten tests were divided into two forms, verbal and non-verbal. Below in Table I is a summary description of the arrangement and other details. Attention is drawn to the last column describing mental processes, most of which are said to have high loadings of $g$. (Freeman, F. S., p. 144).
### TABLE I

**ARRANGEMENT AND DESCRIPTION OF THE TESTS**

**A - VERBAL FORM (Vide Appendix-A)**

<table>
<thead>
<tr>
<th>Serial No. of the Test</th>
<th>Name of the Test in Gujarati</th>
<th>Name of the Test in English</th>
<th>No. of Items in the Test</th>
<th>Mental Functions Tested</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>દિલદારના માનસિક</td>
<td>Reasoning Problems</td>
<td>15</td>
<td>Verbal Analysis of a situation.</td>
</tr>
<tr>
<td>2</td>
<td>કેચુ અને સિમાનત</td>
<td>Verbal Absurdities</td>
<td>27</td>
<td>Using life experiences to detect absurdities.</td>
</tr>
<tr>
<td>3</td>
<td>હજુ દારી સાઝાંત</td>
<td>Analogies</td>
<td>35</td>
<td>Educating relationships or recognizing and utilizing relationships.</td>
</tr>
<tr>
<td>4</td>
<td>બીચ અને સખ્ખાંનત</td>
<td>Matching Professions and Things</td>
<td>30</td>
<td>Acquiring and using vocabulary.</td>
</tr>
<tr>
<td>5</td>
<td>કલાકર કલો બારીમાં</td>
<td>Story Completion</td>
<td>25</td>
<td>Verbal comprehension and arranging ideas in a logical sequence.</td>
</tr>
</tbody>
</table>

**TOTAL:** -- -- 132 --
TABLE I (CONTD.)

B - NON-VERBAL FORM (Vide Appendix-B)

<table>
<thead>
<tr>
<th>Serial No. of the Test</th>
<th>Name of the Test in Gujarati</th>
<th>Name of the Test in English</th>
<th>No. of Items in the Test</th>
<th>Mental Functions Tested</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>મળાડ લડાક્ષ</td>
<td>Matching the legs of the Table</td>
<td>11</td>
<td>Form perception and association.</td>
</tr>
<tr>
<td>2</td>
<td>જરુરા અને દુર્યોગ</td>
<td>Similarities and Differences</td>
<td>46</td>
<td>Insights into similarities and differences.</td>
</tr>
<tr>
<td>3</td>
<td>સ્થળંધક</td>
<td>Classification</td>
<td>29</td>
<td>Analysis and synthesis of material; recognizing and utilizing relationships.</td>
</tr>
<tr>
<td>4</td>
<td>પ્રકૃતિ કંટ્રોલ</td>
<td>Picture Absurdities</td>
<td>32</td>
<td>Using life experiences to detect absurdities.</td>
</tr>
<tr>
<td>5</td>
<td>સ્પેસિયલ</td>
<td>Spatial Relations</td>
<td>32</td>
<td>Visualizing and manipulating objects in space.</td>
</tr>
</tbody>
</table>

TOTAL : 150

====================================================================================================
The Arrangement of the Tests in the Battery:

The arrangement of a group test is either spiral omnibus or discrete. The relative advantages and disadvantages as well as the suitability of using the two methods are discussed elaborately by Desai (Desai, pp. 52-53). He opines that young children in India cannot be relied upon to follow the instructions by themselves. The present lot is younger still and less mature and less experienced with the type of work. It was a foregone conclusion that spiral omnibus arrangement would not be suitable. The method, of course, has its own advantages but nothing good can result from a procedure which is unsuitable. On the contrary, it does more harm than good. Hence the discrete arrangement was chosen, with a full consciousness of its shortcomings.

The Inclusion of the Practice Test:

Desai found a practice test necessary to familiarize the testees with the test material and method of answering as well as to work as a shock-absorber (Desai, p. 75). Hence a short practice test, consisting of four different types of items found in the verbal form was added.

The Undesirability of a Separate Answer Sheet:

A separate answer sheet and reusable booklet is, no doubt, very economical. The problem was to decide whether it was suitable. Cronbach is of the opinion that with young children or persons having little test experience a separate answer sheet may prove confusing (Cronbach, p. 45). Both the conditions are fully applicable to the present group. The
investigator was well aware of the facts and thought that the use of a separate answer sheet would mean sacrificing accuracy and reliability at the altar of economy. Her decision, not to use it, was in accordance with her attempts to simplify the procedure as far as possible.