CHAPTER IV

DESIGN OF THE EXPERIMENT
DESIGN OF THE EXPERIMENT

Introduction:

Reading efficiency, as an independent variable, is a function of two major variables—speed and comprehension. It varies with the change in speed of a reader, as well as change in his comprehension. On an apparent observation it may seem to any man of commonsense that reading speed and comprehension have a linear relationship with each other and the resulting composite score of reading efficiency may vary accordingly. But it is not so. The relationship is more complex than what appears on the surface, for a small increase in speed may result into a disproportionate decrease in comprehension score of the individual and the composite score of reading efficiency may vary in quite an unpredictable manner.

Having realized this phenomenon, the earlier investigators have tried to establish a relationship in the form of standard norms or in the form of a generalized equation between reading efficiency, speed and comprehension. As discussed in chapter III, Elward Fry first tried to establish norms of speed and comprehension for poor, average and good readers. In the Indian context, K. R. Narayanaswamy tried to reform those norms (see Chapter III).
Secondly, measurement of reading efficiency of an individual depends on speed which is expressed universally in terms of words per minute (WPM) and comprehension in terms of comprehension scores (CS). Speed is a continuous variable, like any other variable from natural (physical) sciences, and its measures can be obtained objectively, of course, with the error of reaction time of the subject and the investigator. But the difficulty lies in measuring comprehension and establishing validity, since comprehension itself is a composite variable. It comprises different factors like memory span, length of the material read, reading difficulties in terms of vocabulary and structures, interest in and appeal of the content etc. These factors, however convenient, over-simplify the issue and "under such general headings lies a complexity of interlocking and interactive variables whose relevance one to another is most difficult to establish."\(^1\)

Since the relationship between one factor of comprehension and another is tentative, the weightage of different factors in comprehension cannot be decided.

Comprehension, as an ability, is not just a matter of vocabulary and structures. In addition to these, it is affected by experiential background, intelligence, the family environment and reading habits etc.

Ideally one cannot devise a whole battery of tests measuring each of the constituent elements in the total complex of reading comprehension. This is not feasible. Secondly, satisfactory measurement, requires not only a reliable instrument but also a clear notion of what is to be measured. Hence the investigator was forced to choose some of the factors of comprehension which though they may not cover the entire premises of comprehension would cover at least the major part of it. This could be done only by accepting an operational definition of comprehension.

George Perren suggests: "The first theoretical problem in designing language tests is to agree on a working definition (which may be only a hypothesis of what precisely should be measured)."

Operational Definition of Comprehension:

It was therefore decided to set out to measure the following factors of comprehension while devising tests:

1) ability to get the main ideas;

2. Ibid., pp. 99-100.
ii) ability to get the meanings of essential words and phrases;

iii) ability to find facts;

iv) ability to infer;

v) ability to generalize.

Keeping in view the constructs above comprehension tests were designed and comprehension was sought to be measured objectively. However, the investigator does not claim that the measures of comprehension obtained are very exact and precise since the very nature of comprehension is subjective. At the same time, he does not hesitate to state that a sincere and honest attempt was made to measure this complex variable as objectively as it could best be measured in the circumstances.

**Objectives of the Investigation**

There are three reading speeds (see Chapter III, page No. 113) that a student is likely to require in his adult life for which he should be trained at school and college levels. Of these three, 'average reading speed' is the most useful as it is the speed at which most reading materials are read. It was therefore decided to proceed with the following objectives:

i) to measure the 'average reading' efficiency in English of standard XI students in Aurangabad;
ii) to give students of this class training in silent reading at 'average reading' speed and comprehension;

iii) and to see if this training helped the students to improve their reading efficiency in some way.

**Method of Study**

**Selection of Population**

In order to eliminate the effect of extraneous variables like age, experience, maturity it is necessary to have a homogeneous sample. By selecting students from a particular class the effect of the above mentioned variables affecting the results of an investigation can be minimised to a great extent if not reduced completely. To facilitate proper assessment it was decided to select standard XI, the first year of the Junior college - for such a study. The other considerations in the selection of students of this class were:

1) the students who join this class have to pass the SSC examination which is an external examination and therefore their performance, it was presumed, could be taken as representing a standard;

ii) the English syllabus up to standard X is quite well defined and could guide the investigator to the performance competence that is expected of
the students who pass the examination. This could also help in devising the reading passages;

iii) the syllabus which has been prepared for the Higher Secondary Certificate Examination mentions 'reading comprehension' as one of the objectives. It says:

> The general objectives of both higher and lower level course in English will be to help the students to improve their proficiency in the skills of comprehension (listening and reading) and expression and bring it to a level considerably higher than that aimed at the secondary course.3

iv) to delimit the population, with considerable representation of the entire population, students of Aurangabad city were included. Aurangabad being one of the representative towns in Maharashtra comprises students belonging to:

a) Urban, semi urban and rural areas;

b) different socio-economic status. In addition, both the sexes were represented.

The total number of students in Junior Colleges in Aurangabad city itself exceeds 5000 and this being adequate, the investigator delimited the sample to the Junior Colleges of Aurangabad considering this to be both adequate and representative.

Selection of Passages

To meet the objectives of the study the investigator had to decide as to how many passages were to be selected and what type of passages were to be selected. Taking into consideration the Deogiri experiment it was decided to have in all six passages. The aim in selecting only six passages was:

i) to eliminate factors like mood of the students, fatigue, lack of interest in certain passage or passages, which might affect the results;

ii) a lesser number than six might not probably have had the advantages which are mentioned above and a greater number than this would not have been possible to administer in view of the attitude of the college authorities towards such research projects which they consider to be of little significance in comparison with the regular teaching programme from the examination point of view;

iii) it was also felt by the investigator that a student's rate of improvement can be stably measured by his performance on six different instances;

iv) the students too might have tended to lose interest and would not have taken the tests seriously if the number of passages were more than six;
v) if the number of passages was less than six then any generalization would have meant doing it on the basis of an inadequate data;

vi) and lastly the investigator felt that though six might not be an "ideal" number, it was adequate and could serve the purpose of the study.

No two researchers are likely to agree entirely either on the number of passages or the content of the passages. Though generally an investigator would select passages from standard books which are graded lexically and structurally and where the material is written to rigorous control in terms of vocabulary and structures, one can by no means be certain that the difficulty level is the same. Obviously besides vocabulary and structure there are other factors which make for difficulty in comprehension. There are also no standard norms or criteria regarding the mood, style and nature of the passages. A lot of experimental work therefore had to be done before selecting the passages finally.

Deogiri College Experiment

The first experiment was conducted during September 1977 and February 1978 in Deogiri College of Arts, Science and Commerce. This college was selected with a view to get a group of reasonably representative average students of Standard XI of Aurangabad city. It was thought that among
all the colleges, probably this college was most suitable for the purpose in view, since here he hoped to get students:

i) of all the three faculties;

ii) of different age groups;

iii) of both sexes;

iv) from urban and rural areas;

v) of varying intelligence, experiential background and maturity.

Originally, the plan was to administer ten passages, six passages mentioned below and two each from Edward Fry and K. R. Narayanaswamy. But after administering the first three tests it was realized that the passages from either Edward Fry or K. R. Narayanaswamy were too difficult for these students. They were therefore not included in the course. The college authorities too were reluctant to give so many periods for experimentation. Thus it was decided to administer only six passages which were as follows:

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Name of the passages</th>
<th>Length</th>
<th>Author</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A True Friend</td>
<td>400 words</td>
<td>Raja Nasr</td>
<td>'20 Short Stories' (Longmans Green &amp; Co. Ltd., 1976)</td>
</tr>
<tr>
<td>2</td>
<td>A Girl's Dream</td>
<td>550 words</td>
<td>&quot;</td>
<td>&quot;</td>
</tr>
<tr>
<td>3</td>
<td>Best and Peet</td>
<td>550 words</td>
<td>&quot;</td>
<td>&quot;</td>
</tr>
<tr>
<td>4</td>
<td>The Story of Forty Brothers</td>
<td>730 words</td>
<td>Bhandari</td>
<td>'Graded Short Stories' (Orient Longman Ltd; Reprint 1970)</td>
</tr>
<tr>
<td>5</td>
<td>The Golden Touch</td>
<td>1040 words</td>
<td>&quot;</td>
<td>&quot;</td>
</tr>
<tr>
<td>6</td>
<td>The Grain</td>
<td>1000 words</td>
<td>&quot;</td>
<td>&quot;</td>
</tr>
</tbody>
</table>
The Reading Ease Score of the passages was found to be from easy to very easy. The main consideration in the choice of material was their readability. Discussing the question of how one should decide what is interesting material Narayanaswamy suggests:

The question may be asked: how does one decide what is interesting material? Interest has a psychological as well as a linguistic aspect and while the first one is difficult to determine in the light of agreed, clearly formulated criteria, the second would appear to be less so. A practicable, though hit-or-miss way out of the first difficulty would be to choose material of contemporary value - stories of adventure, travel and exploration, discovery and invention; in short, all that tells of significant things that are going on today and of the people that are engaged in them.1

These also have a special appeal to boys and girls of 15+

age.

The results of the study are summarized in the following table:

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Name of the passage</th>
<th>Reading Base Score</th>
<th>Description of style</th>
<th>No. of students</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A True Friend</td>
<td>88.285</td>
<td>Easy</td>
<td>55</td>
<td>3.36</td>
<td>1.52</td>
</tr>
<tr>
<td>2</td>
<td>Best and Pest</td>
<td>95.007</td>
<td>V.Easy</td>
<td>78</td>
<td>3.36</td>
<td>1.52</td>
</tr>
<tr>
<td>3</td>
<td>A Girl's Dream</td>
<td>93.396</td>
<td>V.Easy</td>
<td>76</td>
<td>4.16</td>
<td>1.53</td>
</tr>
<tr>
<td>4</td>
<td>The Story of Forty Brothers</td>
<td>89.440</td>
<td>Easy</td>
<td>68</td>
<td>4.62</td>
<td>1.63</td>
</tr>
<tr>
<td>5</td>
<td>The Grain</td>
<td>91.135</td>
<td>V.Easy</td>
<td>77</td>
<td>4.37</td>
<td>1.65</td>
</tr>
<tr>
<td>6</td>
<td>The Golden Touch</td>
<td>90.034</td>
<td>Easy</td>
<td>54</td>
<td>3.87</td>
<td>1.58</td>
</tr>
</tbody>
</table>

1 Narayanaswamy K. R., Reading Comprehension at the College Level, op. cit, p. 77.
After the experiment the investigator sought guidance from K. R. Narayanaswamy, (through correspondence) who suggested that "the gap between the shortest passage and the longest passage is too wide" and that "we cannot in all fairness expect a reader to step up his speed unless the passage is reasonably long with the essential information more thinly spread." David P. Harris, on the contrary suggests a different approach:

In as much as the test writer will generally wish to include samples of various kinds of material, the individual test passages should be kept brief. On the other hand, there should be sufficient content to yield at least six or seven comprehension items for pretesting. Passages of between 100 and 250 words are about the proper length.5

We however, decided to accept the suggestion of Dr. Narayanaswamy and revised the length of the passages. The length of each passage was such that it was reasonably long with the essential information more thinly spread and also suitable from a practical point of view, so that the administration of the test was completed within a period of 45 to 50 minutes. The interest of the students was also taken into consideration. It was decided to keep the length of the passages between 650 and 700 words.

Another issue that Narayanaswamy raised was about the source of the passages. His objection was that the "reading passages come from two different sources - Raja Naar and Bhandari" and, therefore, "on the face of it the Bhandari material, in terms of vocabulary alone, is more difficult." He suggested that all the passages be got from the same source, and in case the source material was found not to have been written to control, to rewrite it conforming to the 2000 words limit of the General Service List or any other suitable word list. We therefore selected the following three passages from, 'Graded Short Stories' by Bhandari and Ure (Orient Longman Ltd; Reprint 1970) and rewrite them to the 2000 words limit. The length of these passages was also brought to approximately uniform level. The lengths of the three passages finally selected were as follows:

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Passages</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>The Story of North Brothers</td>
<td>660 words</td>
</tr>
<tr>
<td>2.</td>
<td>The Grain</td>
<td>670 words</td>
</tr>
<tr>
<td>3.</td>
<td>The Golden Touch</td>
<td>675 words</td>
</tr>
</tbody>
</table>

For the three other passages it was decided to select passages bearing on such scientific subjects as were closely related to our day-to-day life in which every individual was expected to be interested. It was believed
that this would help not only in creating interest of the reader in these passages but also in adding an element of variety. Narayanaswamy while discussing the provision of readable or interesting materials in a reading course suggests:

Given the interesting material, even failure to adequately control it will not matter much. Conversely, uninteresting material, even when written to the most rigorous control, is unlikely to be rewarding. Interest aids comprehension and speed; it determines what one reads, the way one reads, and whether one reads at all; above all, it evokes effort, and leads to further reading on one's own.6

The titles of the other three passages and their respective lengths are given below:

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Title of the passage</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Earthquakes</td>
<td>690 words</td>
</tr>
<tr>
<td>2.</td>
<td>The Story of Aeroplane</td>
<td>675 words</td>
</tr>
<tr>
<td>3.</td>
<td>Importance of Petroleum</td>
<td>670 words</td>
</tr>
</tbody>
</table>

To ensure that the content was uniform, at any rate comparable, and that all the passages were written to the same degree of control in terms of vocabulary, the vocabulary of these passages was checked with the General Service List of English words.7 This was found to be

within the range, except for the following words:

I- The Story of Forty Brothers:
   (A) flock, fisted, lap, tulips
   (B) camel, goat, sheep, shepherd, wept.

II- The Golden Touch:
   (A) boon, grant, lit, pitcher
   (B) jug, bed-sheet, palace, pillow, roses, status.

III- Earthquakes:
   (A) areas, barracks, lime, pond, vast.
   (B) cement, clerks, constructed, December, eleventh, fifty, godowns, hundreds, instruments, route, twenty.
   (C) adjust, core, created, interior, layer, normal, sensitive, strained, vary.
   (D) crust, epicentre, focus, seismograph, seismologists, tremor.

IV- The Story of Aeroplane:
   (A) feathers, flapping, flight, thrilling.
   (B) December, hundreds, plane, thousands, two.
   (C) amazing, ballons, currents, vary.
V- The Grain:
(A) crutches, reaped, silt, sinful, valley. (5)
(B) hen
(C) consulted, declared, peasant.

VI- The Importance of Petroleum:
(A) calking, derived, dolls, huge, imprisoned, require, scoped, soothing, testament, utensil. (10)
(B) chemist, diesel, Kerosene, petrol, plastic, radio-cases, railway, scooters, textile, tankers, tar, television, tractor, transistor, tyres, vaseline.
(C) constant, constituents, deposits, eternal, evaporate, fuel, furnace, generator, isolate, locomotives, minor, moisture, proportion.
(D) asphalt, crude, crust, jet, petroleum, paraffin, petra, pitch, residue, synthetic. (10)
The words listed above have been classified into four categories. The words in category 'A' are the ones which are not found either in the General Service List of English Words or in the list of passive vocabulary appended to the report of the Study Group 1967. In each passage, however, a negligible number of these words occur i.e. 4, 4, 5, 4, 5 and 10 respectively. Category 'B' contains scientific terminology. These are again very few, that is, 6 in test No. 3 and 10 in test No. 6. The subject matter of these passages necessitated the use of these words, and hence these words were retained.

Category 'B' contains words which again are not found in either of the lists mentioned above but it was assumed that students would be expected to know these words since they would normally have occurred either in the text-books of earlier classes (for example; goat, camel, sheep etc.) or in their day-to-day life (for instance; cement, godown, petrol, plastic etc.). Category 'C' provides words which are found in the list of passive vocabulary appended to the report of the Study Group 1967. It was for these reasons that words in categories 'B' and 'C' were also retained.

Proper nouns such as names of people and places are not on the 2000 - word list nor on the above list of exceptions, but they contribute little to the understanding
of the passages. The structures in the passages, however, could not be verified with reference to their active or passive acquisition by students since as far as our knowledge goes we do not have any fool-proof method of ensuring it at the level at which this investigation is being made.

Readability of Passages

In order to achieve internal readability Reading Ease (RE) score of each passage was computed according to the formula suggested by Rudolph Flesch. Knowing this accounts only for the linguistic factors and not even for all the linguistic factors at that, but only those (like word length and sentence length) which are measurable. Difficulties can result from less measurable linguistic features and very considerably from content which is not measurable at all. Carl B. Smith concurs with this when he states:

The experience of teachers and case studies remind us that numerous other, probably more significant factors determine the true level of difficulty for any given reader. Experience and background with the topic, interest and enthusiasm for the topic (whether performed or stimulated by the teacher), the density of concepts in the text,

the amount of inference the writer requires the reader to apply — all are involved in measuring difficulty. To date there are no formulas that approach these other factors.9

The Flesch formula is as follows:

\[ RE = 206.335 - .846 (WL) - 1.015 (SL) \]

Where:

- \( RE \) = Reading Ease
- \( WL \) = Word length (the average word length in syllables of the passage being considered.)
- \( SL \) = Sentence length (the average sentence length in the words of the passage.)

The Reading Ease Score of the passages was found to be as follows:

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Name of the passage</th>
<th>RE Score</th>
<th>Description of the style</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The Story of the Forty Brothers</td>
<td>88.26</td>
<td>Easy</td>
</tr>
<tr>
<td>2</td>
<td>The Golden Touch</td>
<td>91.39</td>
<td>Very Easy</td>
</tr>
<tr>
<td>3</td>
<td>The Grain</td>
<td>95.49</td>
<td>Very Easy</td>
</tr>
<tr>
<td>4</td>
<td>Earthquakes</td>
<td>72.13</td>
<td>Fairly Easy</td>
</tr>
<tr>
<td>5</td>
<td>The Story of Aeroplane</td>
<td>82.36</td>
<td>Easy</td>
</tr>
<tr>
<td>6</td>
<td>The Importance of Petroleum</td>
<td>64.17</td>
<td>Standard</td>
</tr>
</tbody>
</table>

To achieve external readability great care was taken in printing the passages, as a page which is not readable can obscure and produce visual discomfort and

fatigue. The passages which were administered during the Final try-out were neatly cyclostyled but the passages administered finally were got printed. The best possible paper was used. The type of print was bold enough. It was the one usually used for text-book printing.

Preparation of Comprehension Tests

Reading comprehension is often tested by asking students to translate a passage from the foreign language into the native language. This type of exercise may be useful for the students who not only understand the foreign language but also have great facility of expression in their own language. The students of the Junior colleges can hardly be considered suitable for such a type of test.

Another common type of test for reading comprehension requires the students to answer in the foreign language or questions on the passage or a section or prose. This may appear to be a fairly valid test of reading comprehension but the success of such a test depends on its careful designing and the ability of the examinee in composing sentences accurately in the foreign language. The assessment of the answers is also hard because the skills involved cannot be separated in the answers and each cannot be given due credit.
A third approach is one where examiners set questions in foreign language and the examinee is asked to answer the questions in his own language. This method helps to overcome the defects of the earlier methods and reduces the bias in favour of the student who has facility of expression in his native language and also helps the student who is weak in composing sentences in the foreign language to gain due credit for understanding what he has read. The validity of this type of test again depends on the construction of the questions.

A more valid method of testing reading comprehension is the use of a set of multiple choice comprehension items testing a variety of relationships (main idea, details, inferences etc.) Wilga M. Rivers proposes:

Reading comprehension, and reading comprehension alone, is more validly tested by selection of the correct answer from multiple choice items in the foreign language, on a reading passage. In this case the student shows his ability to comprehend not only the reading passage but also the fine distinctions between the choices offered him.  

The students who cannot express themselves in foreign language or in their own language do not suffer and are not therefore penalised. Again, like any other earlier mentioned

methods of testing comprehension, the validity of the
test very much depends on its careful construction. The
test may be so framed as to test the student's free or
aided recall of fact or his understanding of the passage.
In other words it should help to test the ability to get
behind the facts and draw the necessary inferences or
both.

For this research project it was decided to design
tests with ten four-choice items. These tests were
constructed in such a way that they tested the student's

i) ability to get the main ideas;
ii) ability to get the meanings of essential
    words and phrases;
iii) ability to locate facts;
iv) ability to infer;
v) ability to generalize;

The language of each test was well within the
comprehension of the average student. The questions put
were such that comprehension was not affected either by
vocabulary difficulties or by a poor structural sense.

Criterion - referenced Tests

In the absence of norm-referenced tests we had to
have criterion-referenced tests. The score on norm-
referenced test tells how well the test taker did in
relation to a national sample of test taker whereas
criterion-referenced tests relate to a specific predetermined objective (criterion) or set of objectives. The criterion score does not compare the test taker with any other norm. Instead, it indicates whether or not a student has achieved a certain level of proficiency. Our criterion was the formula devised by Narayanaswamy, according to which it was decided to assess the reading efficiency. (see Chapter III, p. 123)

The tests were cyclostyled for the try-out but for the final administration these tests were got printed. Half of the test papers for each class were printed on coloured papers. In the coloured test-papers the order of items was also changed so as to reduce the element of copying.

Questionnaires

Two types of questionnaires were prepared. One was meant to be given to the teachers of the colleges to help ascertain their academic and professional qualifications, experience, methods they use to teach students in silent reading, the number of graded books in their library and whether they recommended books to their students for extra reading or not.

Another questionnaire was meant for students to ascertain their achievements in general, and in English in particular, at the SSC examination, socio-economic
conditions, age, sex, reading habits and interest and training in silent reading.

The Try Out

In order to get a group of reasonably representative average students of standard XI, the following three colleges were selected for the try-out of the tests:

1- Saraswati Bhawan College of Arts, Science and Commerce

2- Maulana Azad College of Arts, Science and Commerce

3- Vivekananda College of Arts and Commerce.

The number of students testwise was as follows:

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Name of the passage</th>
<th>No. of students</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>The Story of Forty Brothers</td>
<td>205</td>
</tr>
<tr>
<td>2.</td>
<td>The Golden Touch</td>
<td>249</td>
</tr>
<tr>
<td>3.</td>
<td>The Grain</td>
<td>244</td>
</tr>
<tr>
<td>4.</td>
<td>The Story of Aeroplane</td>
<td>188</td>
</tr>
<tr>
<td>5.</td>
<td>Importance of Petroleum</td>
<td>160</td>
</tr>
<tr>
<td>6.</td>
<td>Earthquakes</td>
<td>174</td>
</tr>
</tbody>
</table>

Administration of Tests: Procedure

The reliability of a test is largely dependent on the manner in which the instrument is employed. Such
administration can cause them to compare poorly with the rest of the group because of factors outside the test itself. A perfect test administration would be one that allowed all examinees to perform at their best under identical conditions. While perhaps this ideal can seldom be attained, every effort should certainly be made to come as close to it as possible. The investigator, therefore, took great care at every stage of administration of tests to ensure that the test administration was as 'perfect' as possible. But in spite of his sincere efforts to attain this ideal there were several factors which were beyond the control of the investigator - factors like large classes, time for administration of tests, examination habits of the students, seating arrangement in the classes etc.

Procedure

The classroom tests were preceded by short talks for about ten minutes on various aspects of reading comprehension. The topics covered in these introductory talks were: the importance of developing reading efficiency in English today; utility of silent reading in general; common faults in reading; the causes affecting reading efficiency and how to overcome them; the interrelationship between reading speed and comprehension; importance of comprehension etc.
Having done this students were given the following oral instructions to ensure uniformity in the administration of the passages and tests:

i) You will be given a passage in English for reading silently. Do not start to read till you are given the signal, "Start".

ii) Avoid common faults like lip movement, nodding your head, or moving the index finger, pen or pencil along the line of print.

iii) Read carefully and as rapidly as possible but do not forget that you are reading to achieve good comprehension.

iv) After you finish reading the passage look up at the black-board and write the time you see on the right hand top corner of the passage provided to you. This indicates the time you have taken in reading the passage.

v) After you have finished writing the time, please fold the paper and keep it in your note-book.

vi) After you have read the passage you will be given a test paper. You should write your name, roll number, faculty to which you belong and the time taken by you, on the spaces provided for these. The test would be of multiple choice items. Read
each item carefully and then put a tick (✓) mark against the one which you feel to be the most appropriate answer. Take your own time to answer the test.

Before distributing the printed sheets containing the passage it was ensured that the students had followed the instructions fully well. The procedure was actually demonstrated.

Then followed the distribution of the printed sheet. The students were given a copy of the passage each. With the signal "Start" they started reading the passage. As the students started reading the passage time was shown on the black-board as suggested by Edward Fry.11

Two columns were drawn up, one indicating minutes and the other indicating seconds. The figures were changed every ten seconds so that at a time there was only one set of figures. No sooner did the students finish reading the passage than they looked up at the black-board and noted the time indicated. The passage was then folded. The students were not allowed to read it the second time. During the reading session the investigator and the lecturer from the college moved round the class and watched the students. Those who moved their lips, or moved their fingers, pen or

11 Edward Fry, op. cit, p. 5.
pencil along the line of print were checked. In the
beginning usually there were quite a few who had these
habits but by the time they reached the sixth test a good
number of them became conscious of these defects and tried
to avoid them. One of the colleagues, who always accompanied
the investigator, helped in writing the time on the black-
board.

Before administering the test it was ensured that
the passages were kept away. Then the test papers were
distributed.

Cheating and copying in examination has almost
become a second nature with some students. In order to
check 'copying' the following precautions were taken:

1. Each test was got printed in two colours. The order
   of items in the two tests was arranged differently.
   The test papers of one type were distributed
   alternately so that students sitting together would
   not find it easy to copy.

2. The supervision of the class was very strict. There
   were three invigilators – the investigator, his
   colleague and the subject teacher.

   For the try-out and the final administration of the
   tests, the tests were usually administered twice a week.
   The administration of tests lasted for 45 to 50 minutes.
The administration of the first test required about 60 to 65 minutes but then gradually the students became familiar with the procedure and it was not difficult to finish the administration of a test in 45 minutes.

The scoring was done by the investigator and the following day the students were given their scores in speed, comprehension and reading efficiency. The students were also asked to increase either speed or comprehension depending on their scores. It was gratifying to note that a few maintained record of their performance of the six tests. This practice of immediate confirmation helped to motivate the students.

The following table gives a comprehensive picture of the results of the tryout.

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Name of the passage</th>
<th>No. of students</th>
<th>Mean of comprehension score</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>The Story of Forty Brothers</td>
<td>205</td>
<td>5.91</td>
<td>1.39</td>
</tr>
<tr>
<td>2.</td>
<td>The Golden Touch</td>
<td>249</td>
<td>4.72</td>
<td>1.69</td>
</tr>
<tr>
<td>3.</td>
<td>The Grain</td>
<td>244</td>
<td>4.23</td>
<td>2.08</td>
</tr>
<tr>
<td>4.</td>
<td>The Story of Aeroplane</td>
<td>183</td>
<td>4.22</td>
<td>1.75</td>
</tr>
<tr>
<td>5.</td>
<td>Importance of Petroleum</td>
<td>160</td>
<td>4.04</td>
<td>1.78</td>
</tr>
<tr>
<td>6.</td>
<td>Earthquakes</td>
<td>174</td>
<td>4.46</td>
<td>2.06</td>
</tr>
</tbody>
</table>
Statistical Treatment

The next step was to improve the test items. This was done by finding out the percent of pull towards every distractor. A criterion was decided in consultation with experts in this field to bring about uniformity in the procedure. The criterion was as follows:

1) Distractors which pull less than 10 percent of the population were to be considered as 'poor'.

2) Distractors which pull more than 50 percent of the population were taken as 'too powerful'.

The 'poor' and the 'too powerful' distractors were removed and substituted by the better ones. Having substituted better distractor for the 'poor' and the 'too powerful', the tests were again administered in Maulana Azad College of Arts, Science and Commerce to verify if each of the distractors cleared the norm set by the investigator. The distractors finally included in the tests for final administration satisfied the criterion decided upon by the investigator.

The items were not arranged in order of difficulty level so as not to disturb their sequential order in the passage. However, their difficulty level differed from one another and ranged between 90 percent and 20 percent.
Test - Retest Reliability

Test reliability may be estimated in a number of ways. Obviously the simplest technique would be to retest the same individuals with the same test. If the results of the two administrations turned out to be highly correlated, it could be assumed that the test had stability. The limitation of this method is obvious. If the time interval between the two testing is relatively short, the examinee's memories of their previous responses will make their two performances spuriously consistent and thus lead to an overestimate of test reliability. On the other hand if the time interval is long this will minimize the 'memory factor', the proficiency of the examinee may have undergone a change, producing different responses to the same items and thus a test could be underestimated.

A second method of computing reliability is with the use of parallel forms - that is, with different versions of the same test equivalent in length, difficulty, time limit, formal and all other similar aspects of tests. This method is probably the best but here again practice effect which can affect the reliability cannot be entirely eliminated.

A third method of estimating the reliability of a test consists in giving a single administration of one form of the test and then by dividing the items into two equivalent halves, (usually separating, odd and even numbered items),
obtaining two numbers for each individual. By such "split-half" procedure one obtains two parallel forms which may be compared and a measure of the adequacy of the sampling provided.

Still another method of estimating reliability is that of "rational equivalence". This method is based not on the scores obtained by each person, but on the pass/fail distribution of each item. This is the most familiar procedure developed by Kuder and Richardson. Here again, as in the case of "split-half" methods, reliability is estimated from a single administration of one form of the test. This method is suitable when our purpose is to verify the internal consistency of the test items.

The coefficients of reliability of the comprehension tests were first calculated by the method of "rational equivalence". Except the first test all other tests yielded the coefficient of less than .5. Later on it was realized that the method was not suitable for the purpose as the test items were heterogeneous in nature, since the items in each test were based on different aspects of comprehension. It was then decided to employ "test-retest" method for estimating the reliability. David P. Harris suggests a similar approach when he says:

Neither the "split-half" nor "rational equivalence" technique of estimating reliability should be used with speed tests. Since there will be relatively few errors in such tests, an odd-even split, for
instance would yield a spuriously high correlation index - one close to "perfect",
in fact. Test-retest or parallel form are the methods best adopted to the measurement
of speed test reliability.12

About 100 students from the sample were randomly selected and retesting was carried out after an interval of fifteen days, believing that this time interval would be adequate to reduce the factor of memory. The reliability coefficients of all the six tests were as follows:

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Name of the passage</th>
<th>Coefficient of reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>The Story of Forty Brothers</td>
<td>.77</td>
</tr>
<tr>
<td>2.</td>
<td>The Golden Touch</td>
<td>.70</td>
</tr>
<tr>
<td>3.</td>
<td>The Grain</td>
<td>.68</td>
</tr>
<tr>
<td>4.</td>
<td>The Story of Aeroplane</td>
<td>.66</td>
</tr>
<tr>
<td>5.</td>
<td>Importance of Petroleum</td>
<td>.60</td>
</tr>
<tr>
<td>6.</td>
<td>Earthquakes</td>
<td>.65</td>
</tr>
</tbody>
</table>

**Validity of Tests**

In the absence of an external criterion either the predictive validity or the concurrent validity of the tests could not be measured. However, it was decided to get the content validity of the tests verified by experts. The opinion of three experts, was sought on the tests and they all approved the content validity of the tests.

Determining Difficulty Level of Passages

To arrange the passages in a suitable order it was necessary to determine their difficulty level. It was decided to ascertain the percentage of passes and failures on each test. 50 percent comprehension score was taken as passing marks. Accordingly the percentages of passes and failures were calculated. Their results are tabulated below:

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Name of the passage</th>
<th>Pass percentage</th>
<th>Fail percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-</td>
<td>The Story of Forty Brothers</td>
<td>71%</td>
<td>29%</td>
</tr>
<tr>
<td>2-</td>
<td>The Golden Touch</td>
<td>63%</td>
<td>37%</td>
</tr>
<tr>
<td>3-</td>
<td>The Grain</td>
<td>30%</td>
<td>70%</td>
</tr>
<tr>
<td>4-</td>
<td>The Story of Aeroplane</td>
<td>35%</td>
<td>65%</td>
</tr>
<tr>
<td>5-</td>
<td>Importance of Petroleum</td>
<td>28%</td>
<td>72%</td>
</tr>
<tr>
<td>6-</td>
<td>Earthquakes</td>
<td>43%</td>
<td>57%</td>
</tr>
</tbody>
</table>

The order of passages for the final administration of test was as follows:

1- The Story of Forty Brothers
2- The Golden Touch
3- Earthquakes
4- The Story of Aeroplane
5- The Grain
6- Importance of Petroleum
The Final Administration of Tests

Population

The research project required the investigator to administer tests in all the Junior Colleges of Aurangabad city. In Aurangabad city there are twelve colleges:

1- Saraswati Bhuvan College of Science
2- Saraswati Bhuvan College of Arts and Commerce
3- A) Maulana Azad College of Arts, Science and Commerce
   B) Maulana Azad College of Arts (Ladies Section)
4- Milind College of Arts
5- Milind College of Science
6- Dr. Ambedkar College of Arts and Commerce
7- Govt. College of Arts and Science
8- Indirabai Mahaliya Mahavidalaya
9- Vivekananda College of Arts and Commerce
10- Vasantrao Naik College
11- Deogiri College of Arts, Science and Commerce
12- Jawaharlal Nehru College of Arts and Commerce

Of these, Saraswati Bhuvan College of Science, Maulana Azad College of Arts, Science and Commerce (A) and Vivekananda College of Arts and Commerce were dropped
at the time of final administration of the tests because these were the very colleges where a tryout was conducted. It was also decided to drop Vasantrao Naik College as it happens to be outside the municipal limits of Aurangabad. Jawaharlal Nehru College of Arts and Commerce was dropped because it is an evening college and most of the students who join this college are employed. They are mature students whose experiential background is usually richer than that of the average junior college student. The Government College of Arts and Science did not allow the investigator to conduct the tests because they were afraid they would not be in a position to complete the syllabus if the tests claimed the students' time. It should be mentioned here that during the year the investigation was carried out (i.e. 1978-79) the colleges started functioning in September and after a month and a half they closed for winter vacation. The final examination had to be conducted before April 1979 as the colleges were to be handed over to the Maharashtra State Board of Higher Secondary Education for Standard XII examination. Standard XI (Science) of Deogiri College also could not take the complete course for similar reasons.

The principals and teachers of some of the colleges where the tests were finally administered were also reluctant in the beginning to spare six periods for such a course and the investigator had difficulty in convincing them of its
utility. The number of students who took the tests, college-wise, is given in Table No. 1.

Procedure of Administration of Tests

The procedure adopted for the administration of the tests was exactly the same as that for the tryout. The administration of tests commenced after the winter vacation during the academic year 1973-74. Tests were administered simultaneously in all the colleges. The timings of the 'periods' varied from college to college. In some colleges periods for English were arranged in the early hours, between 10 a.m. and 11.30 a.m. But in some others they were arranged as late as 3 p.m. The late ones often brought in fatigue and lack of motivation, thus affecting the results to a certain extent. The groups were met usually for two periods of 45 minutes per week. Sometimes there was either only one or no test at all in a week due to co-curricular activities in the colleges.