

**Development of Objective Type Reading Comprehension Achievement Test  
in English for VII Class Students**

**Dr. Madhu Gupta**  
Professor, Dept. of Education, M.D.U Rohtak

**Dr. Jyoti Ahuja**  
Assistant Professor, Vaish College of Education  
M.D.U Rohtak-124001 (Haryana), India

**Abstract:** *This paper has been designed to develop the Reading Comprehension achievement test in English for seventh class students for reading skill. Different steps followed to develop the achievement test were Planning, Preparation, Pre try-out, Try-out, Scoring, Item analysis (difficulty and item discrimination) and Final form of the test. Initially, an objective type achievement test including 100 multiple choice items and fill in the blanks was administered for try-out to 200 students selected randomly from different secondary schools in District Rohtak. Final selection of the items was made on the basis of difficulty value and discrimination index of each item. The investigators selected most of the items of medium difficulty and a few of higher and lower difficulty values were also included. Ebel's (1979) criteria and guidelines were used for categorizing discriminating indices. Split-half method was used for estimation of reliability and 0.90 was found as the calculated value of Reliability coefficient. The test was also validated against the criterion of content validity. It was reported that most of the items were falling in acceptable range of difficulty and discrimination level; however some items were rejected due to their poor discrimination index. In this way, 50 items were selected for final test. The scoring key for the final test was also prepared. The time limit for the final test was 50 minutes.*

**Keywords:** *Item analysis, Difficulty Level, Discrimination Index, Reliability, Validity.*

---

**Introduction**

The main concern of all educational efforts is to see what the learner achieves as achievement is the end product of all educational endeavors. Achievement means accomplishment as proficiency of performance in a given skill or body of knowledge (Good, 1951). Academic achievement holds a cardinal place in the field of education and is the unique responsibility of all educational institution to promote a wholesome scholastic development of the students. It helps the students to understand the hierarchy based on academic achievement. It is an index of success of students' performance, teachers' efforts and significance of curriculum and educational objectives. Academic achievement is related to the acquisition of principles and the capacity to perform efficiently. Assessment of academic performance has been largely confined to the evaluation in terms of information, knowledge and understanding, certain manipulation of objects, symbols and ideas. It is employed as a customary criterion to measure the level of knowledge, understanding and acquisition of skills. Achievement in the educational situation has frequently been referred to as Academic achievement or academic attainment. Dictionary of Psychology (Chaplin, 1965) defined educational or academic achievement as specified level of attainment proficiency in academic work as evaluated by the teacher, by standardized tests or by combination of both. According to Dictionary of Education (Carter, 1959), "Academic achievement means the knowledge attained or skills developed in school subjects, usually determined by test scores or by marks assigned by teachers or both." In the common terminology, academic achievement refers to the level of attainment in various subjects as indicated by marks or grade points after an examination, be it written or oral. It is universally accepted that marks serve the basis of classification and certification, motivation and measurement of educational performance. Academic achievement or academic performance is the outcome of education - the extent to which a student, teacher or institution has achieved their educational goals (Gupta & Lata, 2013). It is commonly measured by examinations or continuous assessment but there is no general agreement on how it is best tested or which aspects are most important -procedural knowledge such as skills or declarative knowledge such as facts (Annie, Howard & Mildred, 1996). Any test that measures the attainments or accomplishments of an individual after a period of training or learning is called an achievement test (Downie, 1984).

Achievement test comprise a very important test in the school evaluation programme and are designed to assess the educational objectives. Such assessment reveals how far the objectives specified have been achieved (Gupta & Lata, 2014). In this situation, when the focus is on the achievement of objectives, teacher's main concern is with the average students to increase the scholastic achievement of the students. Multiple choice questions are the most commonly used tool type for answering the knowledge capabilities of post graduate students in teacher education (Boopathiraj & Chellamani, 2013). Achievement test items are usually distinguished by the kind of response they generate: selected or constructed. The selected response item is often referred to as multiple choices because the test respondent chooses among the choices offered. The constructed response item requires that the respondent generate a written or oral response or a response in the form of a product or process. There is considerable variety in selected and constructed response test items. If item writers are well trained and items are quality assured, it can be a very effective assessment technique. If students are instructed on the way in which the item format works and myths surrounding the tests are corrected, they will perform better on the test (Beckert, Wilkinson & Sainsbury 2003). On many assessments, reliability has been shown to improve with larger numbers of items on a test, and with good sampling and care over case specificity, overall test reliability can be further increased (Steven, 2004). The assessment of student learning requires an adequate and accurate sampling of course content, the multiple-choice test is recommended for measuring achievement at the knowledge, comprehension, application and analysis cognitive levels. Such type of tests often require less time to administer for a given amount of material than would tests requiring written responses. In this way, the investigators decided to develop the objective type reading comprehension achievement test for seventh class students in English for reading skill.

---

## Reading Skill

Reading is one of four language skills that has important role for the students. One's academic success has a strong correlation with reading. One who has a good achievement in classroom usually like to reading. In reading, learners are actively responsible for making sense and catch the ideas of the texts. Reading has different way in interaction with the readers because the writer is not available. That interaction will see in getting all the information from the writer. Brown (1982) says that a person must be able to understand what the author writes in order to be an effective reader. It means that the readers have to catch the information which writer writes so that it can be seen the benefits in doing reading. Academically, reading is one of the most important skills. Reading can be defined as the ability to get understanding from written text. L2 reading can best be understood as a combination of skills and abilities that individuals bring to bear as they begin to read (Grabe, 1991).

## Objective of the Study

The main objective of the present research is to develop an Objective type Reading Comprehension Achievement test in English for seventh class students for reading skill.

## Procedure for Test Development & Data Analysis

To achieve the objectives of the present study, it was planned to develop an achievement test by the investigators in four lessons (**Three Questions, A Gift of Chappals, The Ashes that Made Trees Blom & Quality**) of English Prose of VII grade syllabus. 200 students selected through random sampling technique constituted the sample of the present study. Different steps followed to develop the achievement test Planning, Preparation, Pre try-out, Try-out, Scoring, Item analysis (difficulty and item discrimination) and Final form of the test are written below:

### Planning of the Test

Planning stage of the test tries to answer what content area is to be covered by the test? What types of items are to be included in the test and what are the objectives that are going to be tested? Burton, Brundrett & Jones (2008) observed that the planning stage of a test should include the nature of test and test items and the statement of conditions under which it will be administered. The achievement test was planned with the objectives of measuring achievement in reading skill (English) of VII grade students on selected topics.

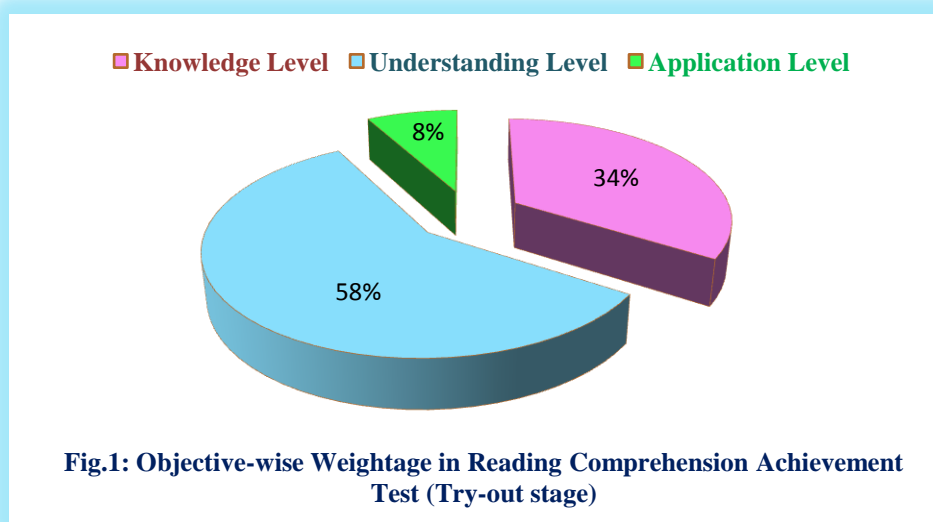
- **Objectives of the test:** For the purpose of constructing achievement test, objectives were defined in behavioral term from selected lessons of Prose. Textbook of class VII prescribed by HBSE Since the major concern here was to test the academic achievement, according to it was decided to test the poor major areas of cognitive domain i.e. knowledge, understanding and application After determining objectives, the learning outcomes were stated as observable terminal performance. In test specifications were developed covering the objectives and subject matter selected to be taught during the experiment.
- **Content of the test:** To decide the weightage to be given to different content areas, objectives and different form of questions, expert opinion of the concerned English teachers was taken into consideration. For the purpose of constructing Achievement Test, objectives were defined in behavioral terms from selected lessons of English textbook of class VII prescribed by HBSE board. Since the major concern here was to test the academic achievement, accordingly, it was decided to test the three areas of cognitive domain, i.e., knowledge, understanding and application. Keeping in view the content a blue print was prepared as shown in given table 1.1. The weightage given to the different instructional objectives is shown through pie chart in Fig.1.1.

Table-1

## Blue Print of Reading comprehension Achievement Test in English (try out)

Contents	Knowledge Objective s	Understanding Objectives	Application Objectives	Total Items	Weightage in %
A. English Reader (Prose) Seen Comprehension	17	34	5	56	56%
B. Unseen (Comprehension)	5	21	1	27	27%
i) Unseen Passage					
ii) Selection of appropriate Title		5		5	5%
iii) Drawing Conclusions		4		4	4%
iv) Vocabulary - Meanings from Context	4			4	4%
v) Word Meanings from Context	4			4	4%
<b>Total</b>	<b>30</b>	<b>64</b>	<b>6</b>	<b>100</b>	<b>100</b>

Note: MCQ = Multiple Choice Questions



**Preparation of the Test**

- **Preparation of the test items:** 100 objective type items fill in the blanks and true/false with wide range of difficulties were constructed from four lessons of Prose in English syllabus prescribed by HBSE for class VII grade. Item were prepared in conforming to blue print. While constructing items it was ensured that no objective remained untested and language of test items was understandable and unambiguous and instructions were clear. The test items were arranged in order of difficulty. The test item were arranged properly and assembled into the test. Easy items was given a place in the beginning and difficult items towards the end. The preliminary draft of achievement test was given to ten expert in education, which include expert in measurement of evaluation experienced english teacher and teacher educators. They were requested to give their opinion about the language and appropriateness of the items. Only those items were selected which were having 80% unanimity. Items that were having difficult language were modified to simple language finally 50 items were constituted the achievement test.
- **Preparation of direction to test item:** Appropriate directions to test items were prepared. The directions were clear and concise, so that the students understand them easily. Test has objective type, true/false and fills in the blanks types questions. Clear instructions were given at the beginning of each section.
- **Preparation of direction for administrations:** A clear and detailed direction as to how the test is to be administered were provided.
- **Preparation of direction for scoring:** To facilitate the objectively in scoring, scoring keys were prepared.

**Preliminary Tryout**

After preparing the test items and scoring key, preliminary draft was administered on a sample of 30 students to find out the ambiguity and adequacy of language. It also helped to detect the omissions or mistake if any, to examine whether the directions to items were actually being followed by students and to examine whether the time allowed was sufficient or not. The problem faced by the students was noted and as a result of preliminary tryout 15 questions are modified.

**Tryout**

The purpose of final tryout was to provide data for determining the discriminating value of item. This also helped to determine the number of item to be included in the final form of the test. The number of the subject in the final tryout was raised to 185. In the preliminary tryout the number of teachers was kept low because the clearing of instructions and the language was to be judged. In the final tryout the number of subjects has to be increased because the investigator had to use the data for item analysis. The achievement test was administered to VII class students individually who have already studied the content. No time limit was fixed for tryout the test. In average students took 60 minutes to answer all the questions.

**Scoring**

After the final tryout the answer sheet were rechecked as per the scoring keys and scoring directions already prepared by researcher one mark was designed to each correct answer and zero to incorrect answer.

### Item Analysis

Item analysis refers to a mixed group of statistics that are computed for each item on a test. The item analysis helps to determine the role of each item with respect to the entire test. The main purpose of item analysis is to improve tests by revising or eliminating ineffective items. There are many different procedures for determining item analysis. The procedure employed in evaluating an item's effectiveness depends to some extent on the researcher's preference and on the purpose of the test. Tabulation is done to determine the following two important characteristics of each item i.e. level of difficulty or item difficulty, and Discriminating power of the test items or item discrimination. The above two indices help in item selection for the final draft of the test. Another step which leads the calculation of item difficulty and item discrimination of a test is item selection based upon the judgment of competent persons as to the suitability of the item for the purposes of the test (Aggarwal, 1986). The following steps were followed for the item analysis.

- All the 185 sheets were arranged in the descending order from highest score scripts at the top to the lowest scores ones at the bottom.
- The 50 upper scripts with highest scores were selected and labeled as "upper group". The 50 scripts with lowest scores were labeled as "lower group" and the middle group of the scripts was set aside. The top 27% of 185 students (50 answer sheets) and bottom 27% were taken into the consideration for computing internal discrimination index and difficulty value. The middle 46% of the total no. of answer sheets were kept aside.

### Difficulty Value

After the formation of two groups, the number of correct responses to an item in each group was marked and tabulated. The difficulty in answering of an item is indicated by the total number of students, who answered it correctly. The larger will be the total number, the easier will be the item. Item difficulty was estimated by determining the percentage of students, who answered the item correctly. The percentage was converted into proportions. The average of the proportions of correct responses on each item in the two end groups was taken to be an estimate of the difficulty value of that particular item.

This point of view is supported by Guilford (1954). Formula for computing the difficulty value 'dv' of each item was:

$$d_v = \frac{P_U + P_L}{2}$$

Where  $d_v$  = difficulty value of the items.

$P_U$  = proportions of correct responses to the items from the upper group.

$P_L$  = proportions of correct responses to the items from the lower group.

### Internal Consistency Discrimination Index (rb)

The relationship between the total scores derived from a test and item scores are referred to as internal consistency discrimination index (rb) of an item. This was found by reading the bi-serial coefficient of correlation between item and total score from the J.C. Flanagan's abac. Flanagan's abac was designed for use, when the sample has been restricted to the highest and lowest 27% of the total score distribution and middle 46% of the examinees of the total score have been eliminated. The proportion of examinees passing the item in the upper criterion group was read on the ordinate and the corresponding proportion from the lower criterion group was read on the abscissa. The value of the coefficient rb was read at the intersection of perpendiculars at these values. When the difficulty values and the internal consistency discrimination indices of each item had been determined, as stated above, items for final draft were selected.

**Item Selection for Final Draft**

Final selection of the items was made on the basis of difficulty value and discrimination index of each item. Lindman (1971) emphasized that easy items should be introduced in a test in order to encourage the students of low ability and some difficult items should be included to challenge the abler students. However, in the interest of constructing a measuring instrument of maximum quality and utility, items were selected whose difficulty value lies in the range **.30 to .79**. The investigators selected most of the items of medium difficulty and a few of higher and lower difficulty values were also included.

**Table-2**

**The Distribution of Difficulty Value (dv) and Internal Consistency Discrimination Index (rb) of All the Items**

<b>Difficulty Value (dv)</b>	<b>F</b>	<b>Discrimination Indices (rb)</b>	<b>F</b>
<b>.00-.09</b>	<b>1</b>	<b>.00-.09</b>	<b>20</b>
<b>.10-.19</b>	<b>3</b>	<b>.10-.19</b>	<b>30</b>
<b>.20-.29</b>	<b>2</b>	<b>.20-.29</b>	<b>3</b>
<b>.30-.39</b>	<b>11</b>	<b>.30-.39</b>	<b>6</b>
<b>.40-.49</b>	<b>10</b>	<b>.40-.49</b>	<b>16</b>
<b>.50-.59</b>	<b>17</b>	<b>.50-.59</b>	<b>11</b>
<b>.60-.69</b>	<b>22</b>	<b>.60-.69</b>	<b>12</b>
<b>.70-.79</b>	<b>19</b>	<b>.70-.79</b>	<b>2</b>
<b>.80-.89</b>	<b>9</b>	<b>.80-.89</b>	<b>--</b>
<b>.90-.99</b>	<b>6</b>	<b>.90-.99</b>	<b>--</b>
	<b>100</b>		<b>100</b>

Garrett (1967) regarded those items satisfactory which are having validity indices of 0.20 or more. Thorndike (1955) considered an item with a validity co-efficient as high as 0.25 as an outstanding 'valid' item. Gronlund (1988) stated, "Zero discrimination power (0.00) is obtained, when an equal number of students in each group answer correctly. Negative Discrimination power is obtained, when more students in lower group answer correctly than the students in the upper group. Both types of items should be removed from norm-referenced tests. The distribution of the all items according to dv and rb has been given in Table 2. Ebel's (1979) criteria and guidelines for categorizing discriminating indices is a widely quoted set of guidelines given in Table 3 was used in this test analysis. Based on the Ebel's guidelines in the table, the 100 test items were categorized as shown in the Table 4. Bivariate scatter diagram for Reading Comprehension achievement test in English between difficulty value (dv) and discrimination index (rb) has been shown in Table 5. The table 5 depicts that the items having dv and rb below 0.19 are rejected and above 0.19 have been retained for final achievement test. Hence, 50 items have been retained for the final form of the Reading Comprehension achievement test in English.

**Table 3**  
**Ebel's Guidelines (Discriminating Powers)**

<b>Discriminating powers</b>	<b>Description</b>
.40 and above	The item is functioning quite satisfactorily
Between 0.30-0.39	Little or no revision is required
.Between .20-0.29	The item is marginal and needs revision
<.19	The item should be eliminated or completely revised

**Table 4**  
**Distribution of Discrimination Powers of the all Items of Achievement Test**

<b>Discriminating Power</b>	<b>Frequency</b>	<b>Item Numbers</b>	<b>Remarks</b>
<b>.40 and above</b>	41	8,9,10,13,22,23,24,25,26,27,36,37,38 , 39,41,50,51,52,56,57,60,62,63,66,69, 70,71,74,75,76,77,78,79,86,87,88,89, 90,98,99,100	Very Good Items
<b>.30-.39</b>	6	11,12,53,55,65,73	Reasonably Good
<b>.20-.29</b>	3	14,40,42	Needs Improvement
<b>&lt; .19</b>	50	1,2,3,4,5,6,7,15,16,17,18,19,20,21,29 , 30,31,32,33,34,35,43,44,45,46,47,48, 49,58,58,61,64,67,68,72,80,81,82,83, 84,85,91,92,93,94,95,96,97	Very Poor Items
<b>Total</b>	<b>100</b>		

Table 5

**Bi-variate Scatter Diagram of Items for Reading Comprehension Achievement Test in English between Difficulty Value (dv) and Discrimination Index (rb)**

Dv Rb	.00-.09	.10-.19	.20-.29	.30-.39	.40-.49	.50-.59	.60-.69	.70-.79	.80-.89	.90-.99
.00-.09	47	54			59,67	19,28,3 1	29,43,4 5	94,96, 64,7	72,30,3 3, 48	18,2
.10-.19		61,8 1	68,8 4	34,49 , 58	35,21	44,95,2 0	91,92,9 3, 85	32,5,4 6, 8, 2,1	80,83,9 7, 3,6	17,15, 16,4
.20-.29						42	14,40			
.30-.39				65,73		53	11,55	12		
.40-.49				25,39 , 69	88,10, 50	52,62	56,57,7 0, 74	9,41,7 6, 86		
.50-.59				66	51	63,99,2 3	75,78, 98, 8	79,13		
.60-.69				26,90	27,36	60,87	37,77,8 9	24,71, 100		
.70-.79						22,38				
.80-.89										
.90-.99										

### Reliability

The investigator has used split-half methodology for estimation of reliability. The items in a test have been split into two tests that are equivalent in content and difficulty. It has been done by splitting among odd and even numbered items. The investigators have found 0.90 as the calculated value of reliability coefficient, which means 90% of the variance of test scores is true-score variance, and only 10% error variance. So, we can say that achievement test is fairly reliable.

### Validity

The purpose of the present investigation and the nature of the test items restricted the use of very exhaustive statistical techniques to validate the test. The test was validated against the criterion of Content Validity which is concerned with the adequacy of sampling of a specified universe of content. To determine content validity the test items and a list of outcomes were given to the panel consisting of five experts in subject matter and three experts in test items. The panel was asked to identify which test item corresponded to which outcomes. The experts agreed with the researcher on the assignment of test items to objectives 95% of time. The percentage was taken as evidence of content validity.

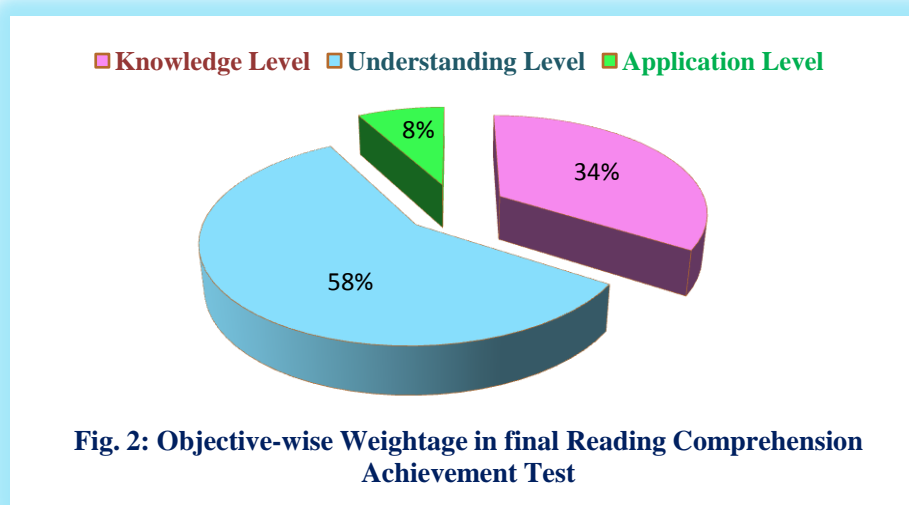
### Final Form of Test

After the selection of items for final test, items were rearranged. On the cover page of the test, directions were printed. The scoring key for the final test was also prepared and has been given along with the final form of the Reading Comprehension Achievement Test, which contained 50 items. The time limit for the final test was 50 minutes. Number of items retained in the final draft of Reading comprehension Achievement Test table 1.6 and the weightage given to the different instructional objectives is shown through Pie Chart in fig. 2.

**Table 6**

**Number of Items Retained in the Final Draft of Reading comprehension Achievement Test at different Cognitive Levels of Objectives**

Cognitive Levels of Objectives	Serial Number of Items Retained	Total
<i>Knowledge Level</i>	9,23,39,40,41,50,53,55,56,66,69,70,71,73, 79, 98,99	17
<i>Understanding Level</i>	8,10,11,12,13,22,24,25,26,36,37,38,51,52, 57,60,62,63,65,74,75,76,77,78,86,87,88,89, 90	29
<i>Application Level</i>	14,27,42,100	4
	<b>Total</b>	<b>50</b>



### Conclusion

The findings of this paper have significance for student teachers and test developer. They should be very careful while selecting items. The size of an acceptable item will depend upon the length of the test, the range of difficulty indices and the purposes for which the test has been designed. The poor items were removed or improved for inclusion in the final test. This work can be repeated in other subjects to develop a good item bank for student community. The principle function of an instrument used in any educational research is to infer student's capacities and it offers information on which to base the making of correct decisions. Developing and administering Multiple Choice Questions on the content knowledge of research methods in education helps teacher educators in molding future teachers. Hitherto item analysis is an important phase in the development of a test or instrument.

---

**REFERENCES**

- Agarwal. Y.P (1986). Statistical Methods, Concepts, Applications and Computations. New Delhi: Sterling Publication.
- Annie Ward, Howard W. Stoker, Mildred Murray-Ward (1996), "Achievement and ability tests - definition of the domain", *Educational Measurement*, 2, University Press of America, pp. 2–5, Available at: <http://nuevaschool.org/~debbie/library/outserch/gender.html>.
- Beckert, L., Wilkinson, T. J., & Sainsbury, R. (2003). A needs-based study and examination skills course improves students' performance *Medical Education* 37 (5), 424–428. [doi:10.1046/j.1365-2923.2003.01499.x](https://doi.org/10.1046/j.1365-2923.2003.01499.x)
- Boopathiraj, C. & Chellamani, K. (2013). Analysis of test items on difficulty level and discrimination index in the test for research in education. *International Journal Of Social Science & Interdisciplinary Research*, ISSN 2277 3630, 2 (2), 189-193. Online available at indianresearchjournals.com.
- Brown, A. D, (1982). Reading Diagnosis & Remediation. Englewood Cliffs New Jersey: Prentice-Hall, Inc.
- Burton Neil, Brundrett Mark & Jones Marion.(2008). Doing Your Education Research Project. UK: SagePublication.
- Carter, C. 1959. Dictionary of Education, pp. 69.
- Chaplin, J.P. 1965. Dictionary of Psychology, pp. 8
- Downie, N.M & Heath, R.W.(1984). Basic Statistical Methods, 5<sup>th</sup> ed., Harper & Row, New York.
- Ebel, R. L. & Frisbie, D. A. (1986). Essentials of education measurement. Englewood Cliffs, NJ: PrenticeHall.
- Garret, H.E. (1967). *Statistics in Psychology and Education*. Bombay: vakil peffer & simons pvt. Ltd; p.189.
- Good, C. V., Ed. (1951). Dictionary of Education. McGraw- Hill Book Co., New York.
- Good, C.V. (1959). Dictionary of Education, Mc-Graw Hill, New York.
- Grabe, W. (1991). Current developments in second language reading research. *TESOL*, 25 (3), 375-406.
- Gronlund, N. E. (1990). Measurement and evaluation in teaching (6<sup>th</sup> ed) New York: Macmillan
- Guilford, J.P. (1954). *Psychometric Methods* (2<sup>nd</sup> ed.). New Delhi: Tata McGraw-Hill co. Ltd; p. 280
- Gupta, Madhu & Lata, Parvesh (2013). Academic anxiety, adjustment and reading interest of male and female students at secondary level: a comparative study. *Global Education Society and Development Journal* 4 (4) 40-48.
- Gupta, Madhu & Lata, Parvesh (2014). Development of Objective Type Achievement Test In Science(Biology) For X Class Students' *International Journal of Social Science and Interdisciplinary Research*, 3(8) ,pp.49-62 .
- Gupta, M & Kapoor, M. (2012). School environment as a determinant of scholastic achievement in English of high school students, *The Educand: Journal of Humanities and Social sciences*, 1(2), 164-170.
- Lindeman, R. H. (1971). *Education Measurement*. Bombay: D.B. Taraporevala Sons & co. pvt. Ltd; p.390.
- Steven M Downing (2004). Reliability: on the reproducibility of assessment data *Medical Education* 38 (9), 1006–1012. [doi:10.1111/j.1365-2929.2004.01932.x](https://doi.org/10.1111/j.1365-2929.2004.01932.x)
- Thorndike, R.L.( 1955). *Personal Selection Tests & Measurement Technique*. New York: John Wiley & Sons, p. 440.