*International Journal of Current Research and Modern Education (IJCRME)*
*ISSN (Online): 2455 - 5428*
*(www.rdmodernresearch.com) Volume I, Issue II, 2016*

# CONSTRUCTION AND STANDARDIZATION OF AN ACHIEVEMENT TEST IN ENGLISH GRAMMAR

**Naresh Kumar**
Assistant Professor, Lala Lajpat Rai Memorial College of Education,
V.P.O. Dhudike, Tehsil & District Moga, Punjab

**Abstract:**

*The paper presents the details of the construction and standardization of the test of achievement in English grammar for IX class under Punjab School Education Board, Mohali. The test initially consisted of 122 items. In order to improve the items, comments of the experts were obtained along with their suggestions and opinions. All the experts positively commented on all the items. So, after the opinions obtained from experts, test contained 122 items in first draft. After the item analysis the final draft was reduced to 50 items. The reliability of the test through split half method was found to be 0.84. In this test concurrent validity was estimated by the investigator. The achievement test scores were correlated with previous examination scores obtained by the students in English subject. The coefficient of correlation was calculated between the achievement test scores and the scores obtained by students in previous examination. The value of 'r' was found to be 0.82, which is very high and the test is regarded as valid.*

**Key Words:** Construction, Standardization, Reliability, Validity, Achievement Test & English Grammar

**Introduction:**

The achievement test is regarded as controlling process for teaching -learning activities. It helps in evaluating the effectiveness of teaching instructions. It also provides the feedback to the students as well as to the teachers. In an achievement test the main emphasis is given on content coverage or course. The achievement test has the focus on the realization of objectives of teaching and learning. The teaching objectives are assessed in terms of terminal behaviours of the students.

The number of achievement tests have been constructed and standardized by various educationists in different fields to measure the academic level of the students. The investigator scrutinized and critically reviewed the existing stock available for testing achievement of IX class students in English grammar. The investigator did not find any appropriate test as per the content part and objectives taken for the present study. Hence it was decided to construct and standardize an achievement test on the selected topics of English grammar for IX class from the syllabus of Punjab School Education Board (PSEB), Mohali.

**Construction of the Achievement Test:**

Construction of the achievement test was carried out in three phases:

- ✓ Planning Phase
- ✓ Construction Phase
- ✓ Standardization Phase

**Planning Phase:**

The following points were taken into consideration during the planning phase:

- ✓ Whom to administer?
- ✓ What to measure?
- ✓ Stage of test administration.

On the basis of above mentioned aspects, it was decided that-

- ✓ An achievement test was meant for students of age groups 14-16 years studying in class IX in different Government Senior Secondary Schools of Moga District of Punjab.
- ✓ Achievement in English grammar was measured. The test aimed at measuring the Knowledge, Understanding and Application aspects of the students in English grammar.
- ✓ Achievement test was administered before the conduct of the experiment (Pre-test), after the completion of experiment (Post-test) and to check the retention of the students after an interval of thirty days after the administration of post test.

Keeping in mind the above said aspects, planning for construction and standardization of the achievement test was done. Further it involved the following two steps:

- ✓ Identification of Contents
- ✓ Identification of Objectives

**Identification of Contents:**

For the present achievement test, the contents were taken from English grammar book prescribed by Punjab School Education Board, Mohali for class IX. To delimit the area of research, the investigator had taken only 12 topics of English grammar. Through discussion with subject experts who had been teaching English subject to class IX students also helped the investigator in identifying concepts. The detail of the content part is given below:

- ✓ The Sentence
- ✓ Subject and Predicate
- ✓ The Noun
- ✓ The Pronoun
- ✓ The Adjective
- ✓ The Verb
- ✓ The Adverb
- ✓ The Conjunctions
- ✓ The Interjections
- ✓ The Articles
- ✓ The Prepositions
- ✓ Modals

**Identification of Objectives:**

Keeping in mind the contents, objectives were identified by the investigator for the construction of the achievement test in English grammar for IX class. Objectives were mainly meant for to test the Knowledge, Understanding and Application aspects of the students.

**Construction Phase:**

The construction of the test was carried out with adequate planning. A blue print was prepared from all the selected topics of English grammar of IX class. Bloom's Taxonomy Objectives formed the basis for the blue print of the test. The blue print of the test was developed in which topic wise weightage was given to each category of educational objectives and number of items for each category of objectives. Topic wise number of items has been shown in the table no. 1.

Table1: Taxonomic Categories and Unit-wise Breakup of Items of Preliminary Draft

| S. No. | ↓ Domain Topic → | Knowledge | Understanding | Application | Total |
|---|---|---|---|---|---|
| 1 | The Sentence | 1, 6, 8, | 2, 3, 4, 5, 7, 9, 12, 13 | 10, 11 | 13 |

| | | | | | |
|---|---|---|---|---|---|
| 2 | Subject and Predicate | 14 | 15, 17, 18 | 16 | 5 |
| 3 | The Noun | 19, 20 | 21, 22, 23, 26 | 24, 25 | 8 |
| 4 | The Pronoun | 27 | 28, 29, 30, 31, 33 | 32, 34 | 8 |
| 5 | The Adjective | 35 | 36, 37, 38, 39, 40, 41 | 42, 43 | 9 |
| 6 | The Verb | 44, 45 | 46, 47, 48, 49, 50, 51, 52, | 53 | 10 |
| 7 | The Adverb | 54 | 55, 56, 59,61, 62 | 57, 58, 60 | 9 |
| 8 | The Preposition | 63 | 64, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78 | 65, 66 | 16 |
| 9 | The Articles | 79, 82, 84, 85 | 80, 81, 83, 86, 87, 88, 89, 90, 91, | 0 | 13 |
| 10 | The Conjunction | 92 | 93, 94, 95, 96, 98, 99, 101 | 97, 100 | 10 |
| 11 | The Interjection | 102 | 103, 104 | 0 | 03 |
| 12 | Modals | 106, 122 | 105, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, | 119, 120, 121 | 18 |
| Total | | 20 | 82 | 20 | 122 |
| Percentage (%) | | 16.4% | 67.2% | 16.4% | 100% |

**Types of Items in the Test:**

Total 122 items were prepared in the preliminary draft of the test. All the items were multiple-choice items, true/false, fill in the blanks, matching type, analogy type, completion type etc.

**Pilot Survey:**

Before individual try out, it was thought essential to have a pre try-out of the test. For this purpose the investigator prepared ten copies of the test and was given to the students to check the language ambiguity and difficulty level of the test. All the students were able to complete the test.

**Preliminary Draft:**

Preliminary draft is necessary in the construction of the test in its final form because in its absence it will be impossible to know how good the test is. While making the preliminary draft of the test, the existing tests in the concerned areas were consulted. New and original test items were prepared to assess the desirable objectives of the blue print. 122 test items of suitable difficulty were complied. Rough idea of the difficulty of the items was obtained by trying out the items on a small group of students as discussed in pilot survey. The items were then edited along with carefully worded instructions, which indicated briefly the nature and purpose of the test. The final manuscript of the preliminary draft was then submitted to different experts for their opinions and criticism, especially to those who have experience of tool standardization and who were teaching English to IX class in Government Secondary Schools of Moga District of Punjab. The division of experts have been given in the table no. 2.

Table 2: Division of Experts who finalized the Final Manuscript of the Preliminary Draft

| S. No. | Category | No. of Experts |
|---|---|---|
| 1 | Teacher Educators | 05 |
| 2 | Teachers Teaching English | 05 |
| **Total** | | **10** |

In order to improve the items, comments of the experts were obtained along with their suggestions and opinions. All the experts positively commented on all the items. So, after the opinions obtained from experts, test contained 122 items in first draft. The investigator personally visited the schools for the administration of

preliminary draft of the achievement test. Before the administration of the preliminary draft proper arrangement of sitting was made, instructions were given carefully and way of giving answer was demonstrated at blackboard. Time factor was also considered by the investigator. Various problems faced by the students were noted by the investigator.

**Formulation of Instructions and Initial Information Regarding Questionnaire:**

The entire set of questions and initial information regarding questionnaire was prepared for the convenience of the selected sample. In this information, nature and type of questions were explained. A scoring key in order to find right and wrong answers was also prepared.

**The Tryout:**

At this stage the preliminary draft was administered on a sample of 50 students of IX class selected from two Government Senior Secondary Schools of Moga District (Punjab) affiliated to Punjab School Education Board (PSEB), Mohali. The description of the sample has been given in table no. 3.

Table 3: Description of the Sample for the Administration of Preliminary Draft of the Achievement Test

| S. No. | Name of the School | No. of Students |
|---|---|---|
| 1 | Government Senior Secondary School, Dhudike, Moga | **25** |
| 2 | Government Senior Secondary School, Moga | **25** |
| **Total** | | **50** |

The time limit was generous at the tryout stage. The test was so timed that nearly 95 percent subjects in the sample completed the test.

**Scoring of the Response Sheet:**

After the administration of the test on 50 students of IX class from the above mentioned schools, the response sheets were taken back and scored with the help of scoring key.

**Item Analysis:**

After scoring the test items, item analysis was carried out. Two kinds of information namely item difficulty and discriminating power of items were computed. Item difficulty means the proportion or percentage of the examinees who have answered the item correctly. And item discrimination power means ability of an item on the basis of which the discrimination is made between superiors and inferiors.

In this test the investigator used Kelly's (1939) method to calculate the item difficulty value and discriminating power. Kelley demonstrated that when extreme groups, each consisting of 27% of the total group were used, the ratio of the difference in abilities of the group to the standard error of their difference, that is the degree of uncertainty about the size of real difference was found to be maximum. Kelley (1939) showed that by taking upper and lower groups of 27% of total sample, one could say with the greater confidence that those in upper group were superior in ability measured by the test of those in the lower group.

For calculating the Discriminating Power (D.P.) the following formula was used:

$$D.P. = \frac{R_U - R_L}{0.5\ N}$$

For calculating Difficulty Value (D.V.), the following formula was used:

$$D.V. = \frac{R_U + R_L}{N}$$

Where
$R_U$ = Number of right responses of the upper group.
$R_L$ = Number of right responses of the lower group.
N = Total number of students in both the groups.
The D.V. and D.P. for each item thus computed and have been given in the table no. 4.

Table 4: Difficulty Value (D.V.) and Discriminative Power (D.P.) of Total Items (N=122)

| S. No. | D.V. | Accepted/Rejected | D.P. | Accepted/Rejected |
|---|---|---|---|---|
| 1 | 0.73 | A | -0.08 | R |
| 2 | 0 | R | 0 | R |
| 3 | 0 | R | 0 | R |
| 4 | 0.81 | R | 0.23 | A |
| 5 | 0.69 | A | -0.31 | R |
| 6 | 0.77 | R | 0.15 | R |
| 7 | 0.73 | A | 0.54 | A |
| 8 | 0.85 | R | 0.15 | R |
| 9 | 0.88 | R | 0.23 | A |
| 10 | 0.42 | A | 0.38 | A |
| 11 | 0.85 | R | 0.31 | A |
| 12 | 0.69 | A | -0.62 | R |
| 13 | 0.73 | A | 0.23 | A |
| 14 | 0.69 | A | 0 | R |
| 15 | 0.58 | A | 0.38 | A |
| 16 | 0.69 | A | 0.15 | R |
| 17 | 0.38 | A | 0.15 | R |
| 18 | 0.85 | R | -0.15 | R |
| 19 | 0.35 | A | -0.077 | R |
| 20 | 0.23 | R | -0.31 | R |
| 21 | 0.62 | A | 0.62 | A |
| 22 | 0.62 | A | -0.15 | R |
| 23 | 0.42 | A | 0.54 | A |
| 24 | 0.42 | A | 0.85 | A |
| 25 | 0.69 | A | 0.31 | A |
| 26 | 0.5 | R | 0.077 | R |
| 27 | 0.92 | R | 0.15 | R |
| 28 | 0.73 | A | 0.38 | A |
| 29 | 0.62 | A | -0.31 | R |
| 30 | 0.38 | A | 0.31 | A |
| 31 | 0.46 | A | 0.31 | A |
| 32 | 0.69 | A | 0.62 | A |
| 33 | 0.19 | R | 0.23 | A |
| 34 | 0.58 | A | 0.54 | A |
| 35 | 0.77 | R | 0 | R |
| 36 | 0.62 | A | 0.62 | A |
| 37 | 0.73 | A | 0.54 | A |
| 38 | 0.58 | A | 0.85 | A |
| 39 | 0.42 | A | 0.23 | A |
| 40 | 0.12 | R | 0.23 | A |
| 41 | 0.62 | A | 0.77 | A |
| 42 | 0.88 | R | 0.08 | R |
| 43 | 0.62 | A | 0.77 | A |
| 44 | 0.85 | R | 0.31 | A |
| 45 | 0.58 | A | 0.38 | A |
| 46 | 0.23 | R | 0.15 | R |
| 47 | 0.5 | R | 0.54 | A |
| 48 | 0.5 | R | 0.54 | A |

| | | | | |
|---|---|---|---|---|
| 49 | 0 | R | 0 | R |
| 50 | 0.42 | A | 0.54 | A |
| 51 | 0.69 | A | 0.62 | A |
| 52 | 0.31 | A | 0 | R |
| 53 | 0.69 | A | 0.62 | A |
| 54 | 0.85 | R | 0 | R |
| 55 | 0.42 | A | 0.85 | A |
| 56 | 0.38 | A | 0.77 | A |
| 57 | 0.5 | R | 0.54 | A |
| 58 | 0.54 | A | 0.92 | A |
| 59 | 0.23 | R | -0.46 | R |
| 60 | 0.5 | R | 1 | A |
| 61 | 0.62 | A | 0.77 | A |
| 62 | 0.58 | A | 0.54 | A |
| 63 | 0.65 | A | 0.54 | A |
| 64 | 0.27 | A | 0.23 | A |
| 65 | 0.42 | A | 0.85 | A |
| 66 | 0.58 | A | 0.85 | A |
| 67 | 0.12 | R | -0.23 | R |
| 68 | 0.54 | A | 0.46 | A |
| 69 | 0.65 | A | 0.38 | A |
| 70 | 0.54 | A | 0.15 | R |
| 71 | 0.77 | R | 0.46 | A |
| 72 | 0.12 | R | -0.078 | R |
| 73 | 0.35 | A | -0.54 | R |
| 74 | 0.65 | A | 0.54 | A |
| 75 | 0.62 | A | 0.62 | A |
| 76 | 0.5 | R | 0.69 | A |
| 77 | 0.62 | A | 0.62 | A |
| 78 | 0.77 | A | 0.15 | R |
| 79 | 0.31 | A | 0 | R |
| 80 | 0.35 | A | 0.54 | A |
| 81 | 0.46 | A | 0.62 | A |
| 82 | 0.69 | A | -0.31 | R |
| 83 | 0.31 | A | 0 | R |
| 84 | 0.85 | R | 0.15 | R |
| 85 | 0.58 | A | 0.54 | A |
| 86 | 0.23 | R | 0 | R |
| 87 | 0.31 | A | 0.15 | R |
| 88 | 0.42 | A | 0.54 | A |
| 89 | 0.77 | R | 0.46 | A |
| 90 | 0.54 | A | 0.46 | A |
| 91 | 0.54 | A | 0.62 | A |
| 92 | 0.65 | A | -0.23 | R |
| 93 | 0.23 | R | 0.46 | A |
| 94 | 0.04 | R | 0.08 | R |
| 95 | 0.5 | R | 0.69 | A |
| 96 | 0.58 | A | -0.38 | R |
| 97 | 0.35 | A | 0.08 | R |
| 98 | 0.08 | R | 0.15 | R |
| 99 | 0.23 | R | 0.31 | A |
| 100 | 0.15 | R | 0.31 | A |
| 101 | 0.38 | A | 0.46 | A |
| 102 | 0.58 | A | 0.08 | R |
| 103 | 0.54 | A | 0.46 | A |
| 104 | 0.27 | A | -0.23 | R |

| | | | | |
|---|---|---|---|---|
| 105 | 0.35 | A | 0.23 | A |
| 106 | 0.5 | R | -0.08 | R |
| 107 | 0.5 | R | -0.08 | R |
| 108 | 0.73 | A | -0.08 | R |
| 109 | 0.92 | R | 0.15 | R |
| 110 | 0.77 | R | -0.15 | R |
| 111 | 0.65 | A | 0.08 | R |
| 112 | 0.77 | R | 0.15 | R |
| 113 | 0.69 | A | 0.15 | R |
| 114 | 0.46 | A | 0.46 | A |
| 115 | 0.46 | A | 0.31 | A |
| 116 | 0.23 | R | 0 | R |
| 117 | 0.15 | R | 0 | R |
| 118 | 0.35 | A | 0.69 | A |
| 119 | 0.19 | R | 0.38 | A |
| 120 | 0.35 | A | 0.69 | A |
| 121 | 0.31 | A | -0.46 | R |
| 122 | 0.77 | R | 0 | R |

For the selection of items, the criteria recommended by Ebel (1966) were used. The evaluation criteria for selection of the test items according to inbox of discrimination power has been given in table no. 5.

Table 5: Ebel's (1966) Criteria for Selection of the Test Items on the basis of Discriminative Power (D.P.)

| **Inbox of Discriminative Power** | **Item Evaluation** |
|---|---|
| 0.40 and Above | Very Good Items |
| 0.20 to 0.39 | Good but Marginal Subject to Modification |
| 0.19 and below | Poor Items |

By following the criteria of D.P. given by Ebel (1966), the items which were accepted, modified and rejected from the first draft has been given in the table no. 6.

Table 6: Selection, Modification and Rejection of Items on the Basis of Discriminative Power (D.P.) of First Draft of Achievement Test

| **S. No.** | **D.P.** | **F** | **Test Items** | **Remarks** |
|---|---|---|---|---|
| 1. | 0.40 and Above | 46 | 7, 21, 23, 24, 32, 34, 36, 37, 38, 41, 43, 47, 48, 50, 51, 53, 55, 56, 57, 58, 60, 61, 62, 63, 65, 66, 68, 71, 74, 75, 76, 77, 80, 81, 85, 88, 89, 90, 91, 93, 95, 101, 103, 114, 118, 120 | Accepted |
| 2. | 0.20 to 0.39 | 22 | 4, 9, 10, 11, 13, 15, 25, 28, 30, 31, 33, 39, 40, 44, 45, 64, 69, 99, 100, 105, 115, 119 | Modified |
| 3. | 0.19 and below | 54 | 1, 2, 3, 5, 6, 8, 12, 14, 16, 17, 18, 19, 20, 22, 26, 27, 29, 35, 42, 46, 49, 52, 54, 59, 67, 70, 72, 73, 78, 79, 82, 83, 84, 86, 87, 92,94, 96, 97, 98, 102, 104, 106, 107, 108, 109, 110, 111, 112, 113, 116, 117, 121, 122 | Rejected |

Table-6 reveals that 46 items have discriminating power equal to 0.40 and above; hence these items were selected for final draft. 22 items have discriminating power between 0.20 to 0.39. These items were revised, modified and also included in the final draft. The remaining 54 items have discriminating power below 0.19, hence these items were rejected.

Test items having D.P. between 0.20 to 0.39 were modified in the light of clarity of language, complexity of contents and difficulty level of items.

The evaluation criteria for the selection of the test items according to inbox of difficulty value has been given in the table no. 7.

Table 7: Kelley's (1939) Criteria for Selection of the Test Items on the basis of Difficulty Value (D.V.)

| Inbox of Difficulty Value | Item Evaluation |
|---|---|
| Above 0.75 | Very Easy Items |
| 0.50 to 0.75 | Reasonably Good Items |
| 0.25 to 0.49 | Marginal Items |
| Below 0.25 | Very Difficult Items |

By following the criteria of D.V. given by Kelly (1939), the items which were accepted, modified and rejected from the first draft has been given in the table no. 8.

Table 8: Selection, Modification and Rejection of Items on the Basis of Difficulty Value (D.V.) of First Draft of Achievement Test

| S. No. | D.V. | F | Test Items | Remarks |
|---|---|---|---|---|
| 1. | 0.75 and Above | 17 | 4, 6, 8, 9, 11, 18, 27, 35, 42, 44, 54, 84, 89, 109, 110, 112, 122 | Very Easy Items (Rejected) |
| 2. | 0.50 to 0.75 | 44 | 1, 5, 7, 12, 13, 14, 15, 16, 21, 22, 25, 28, 29, 32, 34, 36, 37, 38, 41, 43, 45, 51, 53, 58, 61, 62, 63, 66, 68, 69, 70, 71, 77, 78, 85, 90, 91, 92, 96, 102, 103, 108, 111, 113, | Reasonably Good Items (Accepted) |
| 3. | 0.25 to 0.49 | 33 | 10, 17, 19, 23, 24, 30, 31, 39, 50, 52, 55, 56, 64, 65, 73, 74, 75, 79, 80, 81, 82, 83, 87, 88, 97, 101, 104, 105, 114, 115, 118, 120, 121, | Marginal Items (Modified) |
| 4. | Below 0.25 | 28 | 2, 3, 20, 26, 33, 40, 46, 47, 48, 49, 57, 59, 60, 67,72, 76, 86, 93, 94, 95, 98, 99, 100, 106, 107, 116, 117, 119, | Very Difficult Items (Rejected) |

Thus the items having difficulty value between 0.75 and above and below 0.25 were rejected. Items having D.V. between 0.50 to 0.75 were accepted as such. The items whose difficulty values were between 0.25 to 0.49 were accepted after revision and modification. On the basis of D.P. and D.V., out of 122 items, 22 items were accepted as such, 28 items were revised and modified and 72 items were rejected. After the revision and modification of the items, 50 items were selected for the final draft of the test.

**Final Draft:**

After the item analysis of the first draft, the final draft was prepared. In this draft accepted items were included as such. Items required modification were modified by the investigator carefully in the light of discrimination power and difficulty value. After the modification these items were also included in the final draft. Hence total 50 items out of 122 items were selected for the final draft. The breakup of the test items for final draft (Unit wise and Bloom's Taxonomic Category wise) have been presented in the table no. 9.

Table 9: Taxonomic Categories and Unit-wise Breakup of Items of Final Draft of Achievement Test

| S. No. | Domain → Topic | Knowledge | Understanding | Application | Total |
|---|---|---|---|---|---|
| 1 | The Sentence | 0 | 7, 13 | 10 | **3** |
| 2 | Subject and Predicate | 0 | 15 | 0 | **1** |
| 3 | The Noun | 0 | 21, 23 | 24, 25 | **4** |
| 4 | The Pronoun | 0 | 28, 30, 31 | 32, 34 | **5** |
| 5 | The Adjective | 0 | 36, 37, 38, 39, 41 | 43 | **6** |
| 6 | The Verb | 45 | 50, 51 | 53 | **4** |
| 7 | The Adverb | 0 | 55, 56, 61, 62 | 58 | **5** |
| 8 | The Preposition | 63 | 64, 68, 69, 74, 75, 77 | 65, 66 | **9** |
| 9 | The Articles | 85 | 80, 81, 88, 90, 91 | 0 | **6** |
| 10 | The Conjunction | 0 | 101 | 0 | **1** |
| 11 | The Interjection | 0 | 103 | 0 | **1** |
| 12 | Modals | 0 | 105, 114, 115, 118, | 120 | **5** |

| | | | | |
|---|---|---|---|---|
| **Total** | **03** | **36** | **11** | **50** |
| **Percentage (%)** | **6%** | **72%** | **22%** | **100%** |

**Standardization of the Achievement Test:**

The standardization of a test includes the reliability and validity of the test. At this stage the final draft was administered on a sample of 100 students of IX class selected from three different Government Senior Secondary Schools of Moga District of Punjab affiliated to Punjab School Education Board (PSEB), Mohali. The description of selected sample has been given in the table no. 10.

Table 10: Description of the Sample for the Standardization of Achievement Test

| S. No. | Name of the School | No. of Students |
|---|---|---|
| 1 | Government Senior Secondary School, Daudhar (Moga) | 50 |
| 2 | Government Senior Secondary School for Boys, Kokri-Kalan (Moga) | 25 |
| 3 | Government Senior Secondary School for Girls, Kokri-Kalan (Moga) | 25 |
| **Total** | | **100** |

**Reliability:**

Reliability refers to the consistency of scores or measurement which is reflected in the reproducibility of the scores. A test is said to be consistent over a given period of time when all the examinees retain their same relative ranks of two separate testing with the same test. In other words reliability tests that to what extent individual differences of scores can be assigned to chance error. In the words of Anastasi (1968) reliability refers to the consistency of scores obtained by the same individual when re-examined with the same test on different occasions or with different sets of equivalent items or under the variable examining conditions. The following four methods are in the use for estimating reliability of a test scores:

- ✓ Test-Retest Method
- ✓ Alternative or Parallel Forms Method
- ✓ Split-Half Method
- ✓ Rational Equivalence Reliability

**1. Test-Retest Method:** It is the simplest method of estimating reliability index of a test scores. In test-retest method the single form of a test is administered twice on the same sample with a reasonable gap. In this way two sets of scores are obtained by administering a test twice. The correlation coefficient is calculated between the two sets of scores as the reliability index. It is known as coefficient of stability or temporal stability which indicates to what extent the individuals retain their relative position as measured in terms of the test scores over a gap of time. High test-retest reliability shows that there is low variable error in the sets of obtained scores and vice-versa.

**2. Alternative or Parallel Forms Method:** It is an improvement over the test-retest method and it is one way of overcoming the problems of memory, practice, carry over and recall factors. When parallel forms of a test can be developed, the correlation between Form A and Form B may be taken as the measure of the self-correlation of the test or reliability index. Under these conditions, the coefficient becomes an index of the equivalence of the two forms of the test. The reliability index depends on the alikeness of two forms of the test. When the two forms are virtually alike, reliability is too high; when they are not sufficiently alike reliability will be too low.

**3. The Split-Half Method of Reliability:** This method of reliability is an improvement over the earlier both methods i.e. Test-Retest Method and Parallel Forms Method. This method of estimating reliability involves both the characteristics of stability and equivalence. Secondly the test is administered once on the target of sample subjects, while in earlier methods two times administration of test is required after a considerable interval or time gap. This method is most appropriate for homogeneous

items of a test. The test is administered once on the sample of subjects. The scoring is done, after splitting test items into two equal parts. Each individual score is obtained into two parts. The common way of dividing the test items is the odd-even method. In one part odd number of items is assigned and second part will involve even number of items. Then scoring is done separately for these two parts even numbers and odd numbers of items. The coefficient of correlation is calculated of two halves of scores. The coefficient of correlation indicates the reliability of the half test. The self-correlation coefficient of whole test is then estimated by using Spearman-Brown Prophecy Formula.

**4. Rational Equivalence Reliability:** This method of reliability was developed by Kuder Richardson (KR). Therefore it is also known as Kuder Richardson reliability. Lee J. Cronbach has called it as coefficient of internal consistency. This method is based on the assumption that all items have the same or equal difficulty value, but not necessarily the same persons solve each item correctly. In this method, the test is administered once on the sample subjects. The test scores are used for calculating or estimating reliability coefficient. The coefficient of reliability is based on general and specific factors of lasting and temporal situations. Kuder has developed formula for estimating true variance rather than error variance of the test scores. To find out the reliability index of the present achievement test, the investigator used the Split-Half method of estimating reliability. The test was administered once on the target sample as given in the table no. 10. The scoring was done after splitting test items into two equal parts. Each individual score was obtained into two parts. The investigator divided the test items on the basis of odd-even method. In one part odd number of items was assigned and in second part even no. of items were assigned. Then scoring was done separately for these two parts of the test (even numbers and odd numbers of items). The coefficient of correlation was calculated between two sets of test scores. The value of coefficient of correlation was found to be 0.73. The calculated value of coefficient of correlation indicates the reliability of the half test. To estimate the reliability index of the whole test, Spearman-Brown Prophecy Formula was used by the investigator as given below:

$$rtt = \frac{2rt/2\ t/2}{1+rt/2\ t/2}$$

Where rtt = Reliability of whole test.
rt/2 t/2 = Reliability of half test or correlation of two haves

The value of coefficient of correlation of the whole test was found to be 0.84. So the reliability of the whole test was 0.84. The description of reliability of the achievement test has been given in the table no. 11.

Table 11: Reliability of the Achievement Test

| Method | Reliability of the Half Test | Reliability of the Whole Test |
|---|---|---|
| **Split-Half Method** | 0.73 | 0.84 |

**Validity:**

The validity of a test means the degree to which the test actually measures which it purports to measure. The validity provides a direct check on how well the test fulfils its function. Many a time it tells us more than the degree to which the test is fulfilling its function. It actually points out what the test is measuring? Only by examining the criterion, together with the validity coefficient of the test can be objectively determine what the test is measuring? It is for this reason that some psychologists prefer to define validity as the extent to which we know what the test measures?
Validity of a test can be established in five ways.

- ✓ Content Validity
- ✓ Construct Validity
- ✓ Criterion-related validity
- ✓ Empirical or Predictive validity
- ✓ Concurrent validity

**1. Content Validity:** It is one way of gathering evidences which will support the idea that a test measures certain characteristics is to make a careful examination of the test taking situation and the test behaviour in and of itself. When this primary emphasis of the study, the concern is for what is known as 'Content Validity'. Content validity includes several types of validity which depend on types of evidences- face validity, objective validity, sampling validity, curricular validity and factorial validity.

**2. Construct Validity:** Construct validity has also been given other names such psychological validity or factorial validity and trait validity. Construct validity means that the test scores are examined in terms of a construct. Anastasi, 1968 defined construct validity as the extent to which the test may be said to measure a theoretical construct or trait or psychological variable.

**3. Criterion-Related Validity:** A test is validated by correlating the test scores with the external criterion performance.

**4. Empirical or Predictive Validity:** Empirical validity (also called statistical or predictive validity) describes how closely scores on a test correspond (correlate) with behaviour as measured in other contexts. Test scores are used to predict future student behaviour and thus validity is called empirical or predictive validity.

**5. Concurrent Validity:** Concurrent validity is demonstrated when a test correlates well with a measure that has been previously validated. The two measures may be for the same construct, but more often used for different, but presumably related constructs. In this test concurrent validity was estimated by the investigator. The achievement test scores were correlated with previous examination scores obtained by the students in English subject. The coefficient of correlation was calculated between the achievement test scores and the scores obtained by students in previous examination. The value of 'r' was found to be 0.82, which is very high and the test is regarded as valid.

**Conclusion:**

The study was carried out to construct and produce a reliable and valid Achievement Test in English grammar. The test was standardized on the sample of 100 students studying in Government Senior Secondary Schools of Moga district of Punjab. Two research questions which centred on validity and reliability were drawn to guide the study. Theoretical and empirical literature related to the study was reviewed. The reliability of the test was determined through split-half method of reliability which was 0.84 and concurrent validity of the test was estimated which is 0.82. Hence, the constructed Achievement Test in English grammar has a high reliability and validity. The test can be used by the teachers to assess student's achievement in English grammar when they have covered the content areas of IX class.

**References:**

1. Alderson, J. C., Clapham, C. & Wall, D. (1995). Language test construction and evaluation. Cambridge: University Press.
2. Anastasi, A. (1968). Psychological testing. London: MacMillan Company.
3. Ebel, R. L. (1966). Measuring educational achievement. Englewood Cliffis, N. J.: Prentice Hall.

4. Good, C. V. (1959). Dictionary of education. New York: McGraw Hill Book Company Inc.
5. Hasan, M. & Khan, S. (2015). Achievement test in English and Mathematics of secondary school students in relation to gender differences. The International Journal of Indian Psychology, 2, (3).
6. Kelly, T. L. (1939). Essentials of scientific behavioural research. Meerut: R. Lall Book Depot.
7. Webster's Dictionary of Education (1989). Spring Field, Massachusetto: G and C Merriain Co.
8. Zeidner, M. (1987). Essay versus multiple-choice type classroom exams: the student perspective. Journal of Educational Research, 80, 352–358. Retrieved from www.jstor.org/stable/27540265.