

## Construction of an Achievement Test for Students at Collegiate Level

Lalhruaitluangi\*  
Nitu Kaur\*\*

### Abstract

*In the present study an achievement test has been constructed and standardized for the purpose of measuring the achievement of first year collegiate students having the subject of Education either as their core course. The achievement test aims to assess their understanding of psychological construct of 'Learning'. The purpose of constructing achievement test was made in accordance to pre-test post-test quasi experimental design of study. The choice of 'learning' construct was made after ensuring the commonality of the topics covered in class XII Education Syllabi, MBSE (Mizoram Board of Secondary Education) and B.A. 1<sup>st</sup> semester Syllabi of Education, Mizoram University. The test can also be used with higher secondary students towards their completion of higher secondary syllabus who are studying Education subject in the stream of Arts. The content validity of the scale was initially established as per the procedure. The difficulty value and discriminatory index of each item in the initial draft of test was calculated by trial of the test on a small sample of collegiate students of B.A. programme having Education either as their major or minor course. After the item analysis the final test consists of 45 multiple choice questions with 11 items at knowledge level, 18 items at understanding level and 16 items at application level. The reliability of the test was found to be very high using both split-half and KR-20 method. Finally, the norms of the test were established on a set of 200 collegiate students studying Education as core course in different semesters of B.A. programme offered at Mizoram University, Aizawl, Mizoram.*

**Keywords:** Construction, Achievement Test, Students, Collegiate Level

---

\*Lalhruaitluangi, Research Scholar, Department of Education, MZU & Assistant Professor, Dept. of Edn., Pachhunga University College, Aizawl. Email: mahruaieducere020@gmail.com

\*\*Dr.Nitu Kaur, Assistant Professor, Department of Education, Mizoram University. Email: nitukaurmzu@mzu.edu.in

### **Importance of Achievement Test**

Achievement tests are of tremendous importance in the field of education. This purposeful activity is often referred to as Achievement Testing which is any procedure or instrument that is used to measure an examinee's attainment of knowledge or skills (Gregory, 2004). Achievement testing is any measurement process or instrument whose purpose is to estimate an examinee's degree of attainment of specified knowledge or skills (Elliott, 2017). It helps in assessing the learners' competency in a particular area of study. The school system across the globe makes a vigorous use of them in determining if there is any gap between the intended and achieved curriculum and if the curriculum has achieved its learning outcomes. It was in the Fifth International Psychology Congress in Rome, Italy, in 1905, Sante de Sanctis (1862–1935), an Italian doctor, psychologist and psychiatrist, and Binet, in collaboration with Simon, presented the first intelligence tests to the international scientific community (Teive et al., 2017). In partnership with Stanford University, the Stanford-Binet Intelligence test was released in 1916 as a descendant of the Binet-Simon scale (Wilson, 2011). The intelligence test was popularly known as IQ test gave way to the origin of achievement tests in school system throughout the United States realizing the reproducible nature of tests (Poll Everywhere Blog, March 26, 2024). In recent years, there has been increasing emphasis on use of group achievement testing such as Stanford Achievement Tests (Sturner, 2009). The achievement testing has turned out to be a regular exercise in assessment and evaluation of students of different levels of education.

### **Need of the Present Study**

In the present study an achievement test has been constructed and standardized for the purpose of measuring the achievement of collegiate students in their understanding of psychological construct of 'Learning'. Hence the collegiate students of 1<sup>st</sup> year having the subject of Education either as their major or minor course were sample of the study. The purpose of measuring achievement test was made in accordance to pre-test post-test quasi experimental design of study. The constructed Achievement Test on Education Subject at collegiate level is based on the content of the module based on 'Learning' topic, also constructed by researchers. The researchers used the achievement test specifically for measuring accomplishment of students at a point of time, finding difficulties faced by students and assessing the understanding specifically to the topic.

### **Objective of the Study**

Based on the requirement of the study following is the objective of the study:

- To construct and standardize an achievement test on a selected topic of Education Subject for the first year Education core students at collegiate level

### **Procedure of Construction of the Achievement test**

For the construction of test following steps were utilized:

- i) **Planning and Designing the test:** For proper planning of the test the investigators made a decision on to whom the achievement test will be done how achievement of students will be measured. The purpose of the test was to measure the achievement of education core students at first year level. Designing of the test also included objectives of the test, content of the test, nature of the test, number of items, type of items, length of test, weightage to objectives and content allotment of time and scoring scheme. The test was planned on the Educational Psychology paper of collegiate level. It was decided that a time limit of 1 hour shall be fixed for test. The format of question was kept as Multiple-Choice Questions (MCQs) where there was only one correct answer out of the four given options. The blue print of the test was designed based on the three cognitive domains of learning namely the Knowledge, Understanding and Application level.
- ii) **Preparation of the test:** The test was then prepared on the 'Learning' Unit of the Educational Psychology paper. The test was meant for all the 1<sup>st</sup> Semester of B.A. (Education) students. A total of 93 MCQ type questions were made wherein 25 questions were from knowledge level, 31 questions were from understanding level and 37 questions were made on application level.
- iii) **Content Validation of the test:** The test was subjected to content validation by experts of educational psychology using Google forms wherein they were asked to mark the items as Highly Appropriate or Appropriate or Inappropriate. Based on experts' suggestion certain items were modified and discarded from the first draft of the test.
- iv) **Administration of the Test as the First tryout:** The test was tried out on a small sample of students of 1<sup>st</sup>, 3<sup>rd</sup> and 5<sup>th</sup> semester of B.A. (Education). The results of the test were subjected to Item analysis on the basis of difficulty value and discriminatory index of each item.
- v) **Preparing final form of test:** The final form of test was prepared after discarding items on the basis of their difficulty value and discriminatory index.
- vi) **Establishing Reliability of Test:** The test was finally tested for its reliability. As Kuder Richardson 20 (KR 20) formula is the most ideal for of reliability test for the MCQ type items it was employed to find the reliability. Also, the split half method was employed for testing the reliability. The whole test was divided into two halves by segregating odd and even items.

Further Spearman -Brown Prophecy formula of  $2r/(1+r)$  was employed to convert the reliability of half test to whole test.

- vii) Establishing Norms of the Test: both the standard scores of stanines and percentile were calculated on a larger sample to establish the two norms for the constructed achievement test.
- viii) For the construction of the present test for collegiate level, it was initially ensured that the topics covered were also part of class XII Education Syllabi with which the students are well acquainted with. For that purpose, Unit 4 of Education syllabi, Class XII, MBSE (Mizoram Board of Secondary Education), 2020 was referred during item construction which was found comparatively similar content wise to B.A. 1<sup>st</sup> semester Syllabi of Education, MZU, 2020. It was found almost all topics were matching hence suitable for pre-posttest format of experimentation. Table 1 presents the comparison between the two contents.

**Table 1**  
**Content Comparison of B.A. 1<sup>st</sup> Semester Syllabi of Education & Class XII, MBSE, Education Syllabi**

Content of B.A. 1 <sup>st</sup> semester Syllabi of Education, MZU	Content of Class XII Education Syllabi (Unit 4)
• Meaning and Nature of Learning	• Learning
• Theories of Learning and their educational implications	• Concept of Learning.
➤ Trial and Error theory of learning	• Theories of Learning- Trial and Error; Conditioning;
➤ Pavlov's Classical Conditioning	• Insight and their educational Implications.
➤ Skinner's Operant Conditioning	• Laws of Learning.
➤ Kohler's Insightful Learning	• Factors affecting learning.
• Thorndike's Law of Learning	• Types of learning.
• Factors affecting Learning	

Ideally for the construction of the achievement test three distinct phases are followed which are identification of instructional objectives based on Bloom's Taxonomy, preparing design and blue print of test and writing the items/ questions of test.

Based on the content provided in Table 1 instructional objectives were prepared. The test comprises of questions from knowledge (K), understanding (U) and application (A) levels of cognitive domain of Bloom's taxonomy. Bloom's Taxonomy is a powerful tool which provides a framework for breaking down complex cognitive tasks (Davis, 2009). If learning objectives are designed with Bloom's taxonomy in mind, then those objectives can be used to construct test ensuring alignment between instruction and assessment and providing validity to evaluation of students' knowledge and skills (Hall, 2015). The constructed test was proposed to comprise of 20 questions from each of the K, U and A level. However, in the first draft of the test more than 20 questions at each level were made. The first draft comprised of 25 items of K level, 31 items of U level and 37 items of A level which was validated for its content by sixteen (16) experts who rated the items as Highly Appropriate or Appropriate or Inappropriate. The second draft based on experts' validity comprised of

20 items each under K, U and A level of cognition. was tested for finding the difficulty and discriminatory indices.

### Scoring Procedure

The initial draft of achievement test comprised of 60 items. Each item is a multiple-choice question with 1 correct response and 3 incorrect responses. Hence the respondent can get a score of either '0' or '1' on each item. Each item carried 1 mark for correct response.

### The Item Analysis Procedure

The item analysis was done on 57 college students 1<sup>st</sup>, 3<sup>rd</sup> and 5<sup>th</sup> semester of BA(Education). The sampled students were grouped into three groups comprising of top 27% of high scorers as upper group (15 students) and 27% of bottom scorers as lower group (15 students) and remaining middle 46% (27 students) comprised of middle group.

The difficulty value (DV) is percentage of sample in the high and low performing group who answered the item correctly. Popular consensus suggests that the best approach is to aim for a mix of difficulties. That is, a few very difficult, some difficult, some moderately difficult, and a few easy. However, the level of difficulty should be consistent with the degree of difficulty of the concepts being assessed (Tobin). The criteria for classification of the DV are as follows:  $DV < 30$  (too difficult items),  $DV$  between 30- 70% (acceptable/average items),  $DV > 70\%$  (too easy items) and  $DV$  between 50-60% (excellent/ideal items) (Kumar et al., 2021). On the other hand, the discriminatory power/index (DI) is the ability of an item to differentiate between groups of higher and lower abilities and ranges between 0 and 1. Generally  $DI < 0.20$  is considered poor, between 0.21–0.24 acceptable, between 0.25–0.34 (good) and  $DI > 0.35$  is excellent (Date et al., 2019). The formula used to find difficulty level is  $DL = \frac{R_u + R_l}{N_u + N_l}$  and that of discriminatory power  $\frac{R_u - R_l}{N_u}$  (or)  $\frac{R_l - R_u}{N_l}$ , where,  $R_u$  = The number sample in the upper group who responded correctly,  $R_l$  = The number sample in the lower group who responded correctly,  $N_u$  = Number of samples in the upper group,  $N_l$  = Number of samples in the lower group.

Based on calculation of difficulty value (between 0.25-0.75) and discriminatory power (0.2 and above) 15 items were rejected viz. 9 items were rejected from knowledge level, 2 from understanding and 4 from application level. Hence after the item analysis the final test consists of 45 multiple choice questions with a blue print of 11 items at knowledge level, 18 items at understanding level and 16 items at application level. For the purpose of scoring each item followed up with 4 options to choose from, where 3 are incorrect and 1 is correct. Each question carries maximum of 1 mark. Hence the maximum score for the final draft of test is '45' and minimum '0'. Table 2 shows the result of item analysis.

**Table 2**  
**Result of Item Analysis**

<b>Serial No.</b>	<b>Item No.</b>	<b>Difficulty Index (0.25-0.75)</b>	<b>Discriminatory Power (0.2 and above)</b>	<b>Status of Item Retention/Rejection</b>
<b>Knowledge Level Items</b>				
1	1	0.65	0.07	<i>Rejected</i>
2	2	0.81	0.33	<i>Rejected</i>
3	3	0.58	0.67	Retained
4	4	0.44	0.27	Retained
5	5	0.29	0.33	Retained
6	6	0.75	0.33	Retained
7	7	0.37	0.47	Retained
8	8	0.42	0.27	Retained
9	9	0.84	1.07	<i>Rejected</i>
10	10	0.88	0.93	<i>Rejected</i>
11	11	0.86	1	<i>Rejected</i>
12	12	0.7	0.87	Retained
13	13	0.26	0.13	<i>Rejected</i>
14	14	0.24	0.2	<i>Rejected</i>
15	15	0.77	1.13	<i>Rejected</i>
16	16	0.39	0.33	Retained
17	17	0.47	0.33	Retained
18	18	0.82	0.27	<i>Rejected</i>
19	19	0.59	0.47	Retained
20	20	0.65	0.33	Retained
<b>Understanding Level Items</b>				
21	21	0.67	0.67	Retained
22	22	0.32	0.53	Retained
23	23	0.28	0.4	Retained
24	24	0.63	-0.07	<i>Rejected</i>
25	25	0.58	0.33	Retained
26	26	0.14	-0.07	<i>Rejected</i>
27	27	0.53	0.6	Retained
28	28	0.42	0.47	Retained
29	29	0.63	0.6	Retained

30	30	0.25	0.53	Retained
31	31	0.4	0.67	Retained
32	32	0.65	0.73	Retained
33	33	0.32	0.4	Retained
34	34	0.26	0.4	Retained
35	35	0.35	0.33	Retained
36	36	0.29	0.2	Retained
37	37	0.46	0.27	Retained
38	38	0.4	0.4	Retained
39	39	0.37	0.33	Retained
40	40	0.26	0.33	Retained
<b>Application Level Items</b>				
41	41	0.58	0.6	Retained
42	42	0.33	0.2	Retained
43	43	0.68	0.33	Retained
44	44	0.37	0.67	Retained
45	45	0.53	0.67	Retained
46	46	0.49	0.6	Retained
47	47	0.58	0.67	Retained
48	48	0.42	0.4	Retained
49	49	0.79	0.53	Retained
50	50	0.46	0.6	Retained
51	51	0.54	0.6	Retained
52	52	0.23	-0.07	<i>Rejected</i>
53	53	0.61	0.33	Retained
54	54	0.44	0.6	Retained
55	55	0.07	-0.07	<i>Rejected</i>
56	56	0.39	0.33	Retained
57	57	0.54	0.27	Retained
58	58	0.29	0.07	<i>Rejected</i>
59	59	0.4	0.6	Retained
60	60	0.28	0.07	<i>Rejected</i>

### Validity of Test

The first draft of the test was given to 16 experts all belonging to the field of Education. The final draft of the test was content validated again by 7 experts.

### Reliability of the Test

The Split-half Method of reliability test was used dividing the whole achievement test was into odd and even parts. The test was administered on a sample of 100 students of college. Using Spearman-Brown formula the reliability was found to be 0.85, showing a very significant reliability. The KR-20 reliability for internal consistency of test was determined and it was found to be 0.82, which is highly reliable index for test reliability.

### Norms of the Test

For the norms of the test the test scores of 200 students were converted to percentile and stanine scores score. Following Table 3.15 describes the Norms of the test derived from sample of 200 college students.

**Table 2**  
**Norms of the Achievement Test**

Stanine Scores	Norms	Percentile	Norms
01-Mar	Below Average	Below 25 <sup>th</sup> Percentile	Below Average
04-Jun	Average	Between 25 <sup>th</sup> & 75 <sup>th</sup> Percentile	Average
07-Sep	Above Average	Above 75 <sup>th</sup> percentile	Above Average

### Conclusion

The constructed achievement test is solely based on Mizo collegiate level students and serves as a very good tool to classify the competency of students in ‘Learning’ construct in Education subject specifically at collegiate entry level. The achievement test is found to assess students at three primary and lower level of Bloom’s taxonomy of cognitive domain i.e. knowledge, understanding and application. The design of achievement tests varies depending on the use of the test and whether the inference from examinees’ scores is intended to be about an absolute or relative level of mastery of specific knowledge and skills (Elliott, 2017). The present test is a meant for a norm-reference testing in that regard and is useful in understanding the entry level understanding of B.A. (Education) students.



## References

- Date, A. P., Borkar, A. S., Badwaik, R. T., Siddiqui, R. A., Shende, T. R., & Dashputra, A.V. (2019). Item analysis as tool to validate multiple choice question bank in pharmacology. *International Journal of Basic & Clinical Pharmacology*, 8(9), 1999-2003.
- Davis, B.G. (2009). *Tools for teaching*, 2nd edition, Jossey-Bass, San Francisco.
- Elliott, S.N. (2017) Achievement Tests in John Stein (Eds.) *Reference Module in Neuroscience and Biobehavioral Psychology*. Elsevier. ISBN 9780128093245, <https://doi.org/10.1016/B978-0-12-809324-5.05457-2>.
- Forehand, M. (2010). *Emerging perspectives on learning, teaching, and technology, global text*, Michael Orey. (Chapter 3). [https://textbookequity.org/Textbooks/Orey\\_Emergin\\_Perspectives\\_Learning.pdf](https://textbookequity.org/Textbooks/Orey_Emergin_Perspectives_Learning.pdf)
- Gregory J. Cizek (2004). Achievement tests, In Charles D. Spielberger (Eds.) *Encyclopedia of Applied Psychology*, pp. 41-46, Elsevier, ISBN 9780126574104, <https://doi.org/10.1016/B0-12-657410-3/00226-9>.
- Hall, M. (2015, January, 30). A guide to Bloom's taxonomy. *The Innovative Instructor Blog*. [A Guide to Bloom's Taxonomy | The Innovative Instructor \(jhu.edu\)](https://www.innovativeinstructor.com/a-guide-to-bloom-s-taxonomy/)'
- Kumar, D., Jaipurkar, R., Shekhar, A., Sikri, G., & Srinivas, V. (2021). Item analysis of multiple choice questions: A quality assurance test for an assessment tool. *Medical Journal Armed Forces India*, 77, S85-S89.
- Raymond Sturner (2009). General principles of psychological testing in William B. Carey, Allen C. Crocker, William L. Coleman, Ellen Roy Elias, Heidi M. Feldman (Eds.) *Developmental-Behavioral Pediatrics* (Fourth Edition), W.B. Saunders, pp. 763-769, ISBN 9781416033707, <https://doi.org/10.1016/B978-1-4160-3370-7.00077-8>.
- Teive, H. A. G., Teive, G. M. G., Dallabrida, N., & Gutierrez, L. (2017). Alfred Binet: Charcot's pupil, a neuro psychologist and a pioneer in intelligence testing. *Arquivos De Neuro-Psiquiatria*, 75(9), 673–675. <https://doi.org/10.1590/0004-282X20170097>
- Tobin, M. A. (n.d.). *Guide to item analysis. institute for teaching excellence*. The Pennsylvania State University. Retrieved from [GuideToItemAnalysis.pdf](https://www.psu.edu/teaching-excellence/guide-to-item-analysis/) (psu.edu) on 19th Feb., 2024.
- Wilson, S.M. (2011). Stanford-Binet intelligence scales. In: Goldstein, S., Naglieri, J.A. (eds) *Encyclopedia of Child Behavior and Development*. Springer, Boston, MA. [https://doi.org/10.1007/978-0-387-79061-9\\_2783](https://doi.org/10.1007/978-0-387-79061-9_2783)
- Retrieved from [Achievement Tests: Definition, Types & Best Practices \(polleverywhere.com\)](https://www.polleverywhere.com/Achievement-Tests-Definition-Types-&Best-Practices.html) on 26<sup>th</sup> August, 2024