Construction and Validation of Science Achievement Test at Secondary Level

Priya Kumari
Research Scholar (SRF)
Faculty of Education, Banaras Hindu University, Varanasi
Uttar Pradesh

Dr. Kishor H. Mane
Assistant Professor
Faculty of Education, Banaras Hindu University, Varanasi

Abstract

Achievement test is an assessment tool of academic outcomes of students in particular subject. The primary goal of an achievement test is to assess the mastery of knowledge in a specific subject or content. This test measures the capability of an individual in quantitative terms that shows how much individual has expertise in that particular subject. Generally achievement test is mostly used by teachers in school and colleges to assess the learning outcomes of individual in teaching learning process at the completion of the unit. This paper represents the construction and authentication of Science Achievement Test (SAT) at secondary level for 9th standard students. The researcher constructed 70 multiple choice items based on NCERT syllabus of Science taken from 3 units of biology section as mentioned "Fundamental Unit of Life: Cell", "Tissue" and "Diversity of Living Organism". After the item analysis 54 items were retained in the final version of tool. Test retest reliability was calculated and value of coefficient of correlation was found to be 0.78 and Cronbach α was 0.87 and validity of the test established by content validity and face validity through experts' suggestions. This test is developed especially to knowing the learning-outcomes of students in science concepts that was taught by mind mapping techniques. The items of the test were based on the selected mind maps topics such as structural organization of cell, Nucleus, Cytoplasm etc. Mind maps are creative and logical way of representing ideas in nonlinear form and showing how you connect ideas and concepts. They allow students to create visual images to make information more meaningful and allow for deep learning to occur. Mind mapping can enhances the

students' achievement and knowledge retention because learner constructing the knowledge via understanding the concepts.

Keywords: Science Achievement test, Reliability, Validity, Item difficulty and Item discrimination.

Introduction:

Achievement: Desired by All, Defined by Few

In the Dictionary of Education, Achievement is defined as (1) accomplishment or proficiency of performance in a given skill or body of knowledge; (2) progress in school (C. V. Good, 1973, p. 7). Academic achievement is defined as "knowledge gained or competencies developed in the school subjects, usually denoted by test scores or by marks assigned by teachers, or by both (p. 7).

Science Achievement Test

"A test of educational achievement is one designed to ensure knowledge, understanding or skills in a specified subjects or a group of subjects." –Freeman mentioned by (Bhatnagar & Bhatnagar, 2014). Science Achievement test is an assessment tool that measure the quality of students' general ability, language achievement, retention and reproduction ability of basic facts, knowledge and skills attained in a given time period. The steps of science test construction were followed as per guideline mentioned in Kachhap (2020) given by Linn & Grondlund (1990). The objective of the science achievement test was to assess the science achievement among students of secondary school. The test items were constructed on the basis of Bloom's classification of cognitive domain (1956) as mentioned knowledge, Understanding and application levels. The lowest level of learning outcomes is knowledge, which is defined as recalling and remembering
 a broad range of content, from particular information to entire theories. The ability to comprehend the material by translating, explaining, and predicting consequences or effects is defined as comprehension/understanding, while the ability to apply what you have learned in new and concrete situations using rules, techniques, ideas, method, concepts, laws, and theories is referred to as application. Linn and Gondlund (2009) demonstrate that these learning outcomes necessitate a higher level of understanding.

Need of the development of Science Achievement test

The researcher analyzed different available online and offline achievement test in science for 9th standard students. After analyzed various available science achievement tests Bhagat & Baliya (2016), Sharma & Sarita (2018), Agrawal (2018) & (Ahmad, Sultana et al. 2020) found that except a few, most of the test were subject and content specific, i.e. biology, physics and chemistry. The researcher's intended to test the learning outcomes of students in selected science concepts which were suitable according to the objective of the study. The Objective of the study was to assess the effectiveness of mind mapping technique on learning outcomes of secondary school students in science, so the researcher constructed the appropriate tool for this purpose. Therefore, researcher develop an achievement test based on the selected topics from 3 chapters as mentioned "Fundamental Unit of Life", "Tissues" and "Diversity in living Organism". These test items were based on

biology topics that were taught to the students in mind map forms. Mind maps are non-linear general interpretations that "comprise a network of interconnected and related themes or any concept/idea can be connected to any other" their purpose is to stimulate associations among ideas or concept, and their formation requires free-form and spontaneous thinking. A mind map is loosely structured around a main idea that is positioned at the centre and grows outward organically in all directions as mentioned by (Buzan, 1996). According to (Davies, 2011), the tool's potential and impacts on student learning are as follows: as a student works on visually representing a complex set of relationships, s/he is more likely to understand those relationships, remember them, and be able to analyze their component parts; it also promotes 'deep' rather than 'surface' approaches to learning.

Objective of the study

To prepare and construct the Science Achievement test for 9th Standard of NCERT Syllabus

Methodology Used

The following steps were considered during the construction of Science Achievement test as mentioned guideline by Linn & Grondlund (1990).

- Planning of the test
- Preparation of the test
- Administration of the Test /Pilot testing/First tryout
- Final tryout of the test
- Preparing final form of test
- Standardization of test: Establishing Reliability &Validity

Planning of the test:

With this aim researcher reviewed various standardized achievement test and develop an idea for construction of the test. The investigator kept following aspect in mind such as: to whom, what, when and how to measure the learning outcomes of students. The researcher analyzed the content of science book of NCERT of class 9th (CBSE Board) and selected the topics from biology units for preparation of achievement test. These units are covered in their first phase of examination. The comprehensive description of selected chapters and subunits has been showed in table.

Table No.1

THE DESCRIPTION OF SELECTED CHAPTERS AND UNITS

SUBJECT	CHAPTERS	SUBUNITS
	Fundamental Unit of Life: Cell	Structural organization of the cell, Plasma membrane, Cell wall, Nucleus, Cytoplasm, Endoplasmic Reticulum, Golgi apparatus, Lysosomes, Mitochondria, Plastids, and Vacuoles
		Plant Tissue, Meristematic Tissue, Permanent Tissue,

| BIOLOGY | Tissues | Simple Permanent Tissue, Epidermis, Complex Permanent tissue, Animal Tissue, Epithelial Tissue, Connective, Muscular, and Nervous Tissue. |
| | Diversity in Living Organisms | Hierarchy of Classification, 5 Kingdom Classification of Whittaker, Plantae, Animalia, Chordata, and Vertebrata. |

Determining the Instructional Objectives of Science Achievement Test

In the construction of Science Achievement test, Instructional objectives are contributing an important role in instructional process. Instructional objectives are clearly indicated the direction of teaching and learning and communicate with students and guiding how to assess the students' learning outcomes as reported by Kachhap (2020) given by (Linn & Gronlund, 1995). Instructional objectives are related to learning as students gain learning experience (Process = study of Tissues) and learning outcomes (Product = Knowledge about tissue, differentiate between plant and animal tissue and functions of tissue) shown by (Linn & Gronlund, 2009).

Construction of Science Achievement test

In this phase designed the specified weightage given to different component of the science achievement test as mentioned:

- Instructional Objectives
- Types of items
- Units and Subunits of the content
- Level of Difficulty

As the NCERT (2005) syllabus showed a large body of research on science education which recommended pedagogy at secondary level should be hands on experiences and inquiry based. These learning experiences are essential for imbibing the spirit of scientific inquiry in an experiential manner of learning science things that are directly related to the child's experience and carrying out activities followed by discussion. NCERT syllabus (2022) also recommended that science subject has crucial role in developing well defined abilities in cognitive, affective and psychomotor domain of students and it also augment the spirit of inquiry, objectivity, creativity and aesthetic senses in students and for ensuring the success of one's teaching educational progress should be measured according to the definite objectives (Mohan, 2002) reported by Kachhap (2020).Therefore, it is important to assigning weightage to various objectives. Instructional objectives serve as the benchmark for multiple mental processes during teaching and learning. For maintaining the objectivity in test construction researcher has decided to find out teaching fraternity and weightage are also given to different levels. Hence, researcher presented the various levels of weightage distribution: weightage to objectives/learning outcomes, weightage to chapters, weightage to units and subunits and weightage to level of difficulty.

Table No.2

WEIGHTAGE TO OBJECTIVES/ LEARNING OUTCOMES

Sr. No.	Objectives	No. of Items	Percentage
1.	Knowledge	35	50%
2.	Understanding	28	40%
3.	Application	07	10%
TOTAL		**70**	**100%**

Table No.3

WEIGHTAGE TO THE CHAPTERS

Sr. .No.	Chapters	No. of Items
1.	Fundamental Unit of Life: Cell	23
2.	Tissues	29
3.	Diversity in Living Organisms	18

Table No.4

WEIGHTAGE TO UNITS AND SUBUNITS

Sr. No.	Units	No. of items	Percentage
1.	The Fundamental Unit of Life: Cell	23	**32.85%**
2.	Plant tissue	14	**20.00%**
3.	Animal Tissue	15	**21.42%**
4.	Classification of Plants	10	**14.28%**
5.	Classification of Animals	08	**11.42%**
	Total	**70**	**100%**

Table No.5

WEIGHTAGE TO LEVEL OF DIFFICULTY

Sr.No.	No. of Items	Level of Difficulty	Percentage Value
1.	28	Easy	**40%**
2.	26	Average	**37%**
3.	16	Difficult	**22%**
Total	**70**		**100%**

Construction of Blueprint of Science Achievement test:

Gronlund and Linn (2000) emphasized that table of specification referred to as subject matter of a course or curriculum that can be well defined to include both particular subject matter and instructional objectives as mentioned by (Ahmad & Sultana, 2020). This step involved in the preparation of blueprint. Blue print is the overall roadmap or summary of content based on the instructional objectives. This phase is emphasized with, how many questions to be prepared for selected content area and their percentage distribution over the different objectives and given weightage accurately. On the basis of blue print 70 items (MCQ)

Multiple choice questions were prepared for selected units of science (Biology). Selected units were taught by the researcher that was selected for items preparation. Items were prepared in bilingual (English and Hindi) for easy understanding to students. While preparing the items language were considered as easy and simple to reader and without any ambiguity. The blueprint has been shown in table. All items were of one mark each. Number of items was 70 and total marks for items were 70.

Table No.6

Blueprint of Science Achievement test

OBJECTIVES	KNOWLEDGE	UNDERSTANDING	APPLICATION	TOTAL
Weightage	**50%)**	**(40%)**	**(10%)**	**(100%)**
The Fundamental Unit of Life	10 (1)	12 (1)	1 (1)	**23**
Plant Tissue	6 (1)	4 (1)	4 (1)	**14**
Animal Tissue	8 (1)	5 (1)	2 (1)	**15**
Classification of Plants	5 (1)	5 (1)	-	**10**
Classification of Animals	6 (1)	2 (1)	-	**8**
TOTAL	**35**	**28**	**7**	**70**

Administration and Selection of Items:

1. Pre-tryout/Individual tryout-

All the test items were evaluated by 10 experts from the field of education, science teachers and language experts. Valuable suggestions were provided by experts related to the appropriateness, language, content, distracters and applicability of items. After finalizing the test items administered to the pre-tryout to test the credibility of test items on a group of 25 students of 9th standard. In this process of administration individual students were noticed by researcher and researchers given the proper instructions how to start the test. The participants were given proper time to response to the items on Science Achievement Test.

2. Final tryout-

This Science Achievement test was given to 100 9th-grade students as a final tryout. After responding the test items were collected from students and check the responses by giving 1 mark for correct answer and 0 mark for wrong answer to students. After scoring the responses of SAT, item analysis was done to check the effectiveness and suitability of the test items to be considered in the test. These question booklets were examined through item wise to find out the   difficulty level and discriminating power of the test items. Information on item difficulty can also help in determining how well students in general have achieved instructional goals and identify the learning needs of individual students (Stodola & Stordahl, 1972). The following procedure was followed by researcher for item analysis:

Step1. After collecting responses of the students, the test papers were arranged in descending order (with high to low scores).

Step2. According to Kelly (1939) as described by Kachhap (2020), in this step the top scorer of 27% and bottom scorer of 27% of the papers were selected and rest of them leave. These 27% papers from top (27) were considered as "Upper Group" and 27% papers from bottom level (27) as "Lower Group".

Step3. Then the responses of the students analyzed one by one and individual students wise all items from upper and lower groups by assigning scores of right and wrong.

Step4. Counted the "discrimination power and difficulty value"

Formula for Discrimination Power:

Discrimination Power = RU-RL/N

Difficulty Value = (RU+RL/2N)*100

Where, RU = No. of students of Upper 27% group

       RL = No. of students of Lower 27% group

       N = No. of students in each group

Item Difficulty:

Item difficulty is referred as the number of people who get a particular item correct. The proportions of item difficulty do not really indicate item "difficulty" but item "easiness". The higher the proportions of the students who get the item correct, the easier the item (Robert & Dennis, 2001). Its index ranges from 0 to +1.00. Item those have higher difficulty indexes show easier items. An item with a difficulty level of.75 is correctly answered by 75% of the examinees. An item with a difficulty level of .35 is correctly answered by 35% of the examinees (McCowan & McCowan, 1999).

In the majority of tests, 0.3 to 0.7 difficulty ranges of items tend to maximize information about individual student differences (Robert & Dennis, 2001).

Discrimination Index:

"The discriminating power of a test item refers to the degree to which it discriminates between pupils with high and low achievement"- Gronlund

Discrimination index describes whether people who performed well on specific items also performed well on the whole test. This method compares students who carried out very well on a test to those who carried out very poorly on a test and then determines the proportion of students in each group who correctly answered each item. The difference between these proportions is called the discrimination index.

Table No.7

Index of Discrimination suggested by Ebel & Frisbie (1991)

Sr. No.	Index of Discrimination	Item Evaluation
1.	0.40 and above	Very Good items
2.	0.30 to 0.39	Reasonably Good items
3.	0.20 to 0.29	Average items
4.	Below 0 to 0.19	Poor items

Table No.8

Item Analysis Final Try Out Draft

Discrimination Power	Range of Items (0.2-0.8)	Remarks	Total
Above 0.40	2,5,7,11,14,15,16,17,20,21,22,27,28,31,32, 33,34,37,38,42,48,57,58,59,60,61,66,67	Very Good items	**28**
0.30-0.39	1,12,13,19,23,24,26,44,54,63,64,65	Reasonably Good items	**12**
0.20-0.29	4,6,8,18,29,39,47,49,50,53,55,56,62,69	Average items	**14**
Below 0.20	3,9,10,25,30,35,36,40,41,43,45,46,51,52,68,70	Poor items	**16**
TOTAL			**70**

Above found scores of index of discrimination of the items, the items having discrimination power range from 0.2 to 0.8 were considered for the final draft of the Science Achievement test (Ebel & Frisbie (1991). The final draft of test contained only 54 items after eliminating these items 3,9,10,25,30,35,36,40,41,43,45,46,51,52,68,70.

Preparation of Final Draft:

After a thorough appraisal of the test items that were based on item analysis. The researcher finalized 54 items in the final draft of the Science Achievement test. The researcher created the final form of the test booklet as well as the item scoring key. The final test was further standardized through experimental validation, establishing reliability and validity.

Reliability and Validity of tool:

Reliability is one of the necessary characteristics of any research tool. Reliability refers to the consistency of measurement, that is, how consistent test scores or others assessment results are from one measurement to other (Linn & Gronlund, 2009). The extent to which test results are consistent, stable, and free of error variance is defined as reliability, which is the most reliable single measure of test accuracy. Coefficient Alpha or KR-20 measures the extent to which a test provides the same ranking of examinees when re-administered. Reliabilities as low as .50 are satisfactory for short tests of 10 to 15 items, but tests with more than 50 items should have reliabilities of .80 or higher. If reliability is less than .80, a single test score should not be used to make correct decisions about individuals (McCowan & McCowan, 1999). In this research work, two forms of reliability: Cronbach's alpha and Guttman Split-Half coefficient methods were used by the researcher through

which the coefficient of reliability was found 0.87 and 0.78 respectively. These values were indicating that the test is reliable.

Validity is important key to effective research. The validity of an instrument refers to whether or not it actually measures what it claims to measure. In other words, rather than being an absolute state, it should be viewed as a matter of degree (Gronlund, 1981). This Science Achievement test was validated using content validity, which means that the instrument must demonstrate that it fairly and comprehensively covers the domain or items that it claims to cover (Carmines and Zeller, 1979, p. 20), as cited in (Cohen, Manio, and Morrison, 2000). Experts in the area of study established the test's content validity. The items of the tool were assessed by the experts for their relevance and also establishing the pre- tryout step. After validating the reliability and validity this science achievement test was prepared for experimental study.

Scoring Key of Science Achievement test:

This research tool contained 54 multiple choice questions. Every item possesses four alternatives in which one was the correct option and rests of all were distracters. For each correct options 1 score was given and for wrong answers 0 score was provided. The maximum marks for total items were 54 and there were no negative points for wrong answers.

Conclusion:

In this paper researcher constructed and validated the science achievement test for assessing the learning outcomes of secondary level school students in science subject. Generally achievement tests have been used for assessing and measuring the student progress or mastery in a particular content area or subject to know how much student have domain specific knowledge of particular subject at the end of course, content, unit or chapters. The items in this tool were based on the topics selected for mind maps as Structural Organization of cell, Plant tissues, Animal tissues, 5 Kingdom Classification, Plantae and Animalia. This tool was standardized on the sample of 100 students of class 9th of CBSE Board School, Varanasi. The reliability of the tool was established by Cronbach's alpha and Guttman Split-Half coefficient methods which was found 0.87 and 0.78 respectively and content validity was determined by field and subject expert. Achievement tests are also used in guidance, counselling, and remedial education programs. Achievement tests can also help teachers assess students' difficulties in the teaching-learning process. It also serves as an evaluating tool for determining the applicability of instructional techniques and teaching methods, as well as determining or deciding on the appropriate curriculum content for students.

References

Anastasi, A. & Susana, U.(2018). *Psychological Testing*, (7th Eds.), Pearson India Education

     Services, Inc,

Agrawal, V. (2018).Construction and Validation of an Achievement Test for Concept

     Assessment in biology. *International Journal of Education*, 9, 32-56.

     http://ijoe.vidyapublications.com

Ahmad, S., Sultana, N. & Jamil, S. (2020). Considerations for Constructing and Validating   Biology
Achievement Test at Secondary Level. *US-China Education Review B*, 10(1),
     13-25. doi: 10.17265/2161- 6248/2020.01.002

Bhagat, P. & Baliya, J.N.(2015).Construction and Validation of Achievement Test in Science.
     International Journal of Science and Research, 5(6), 2277-2280.
     http://dx.doi.org/10.21275/v5i6.NOV164808

Bhatnagar, A.B & Bhatnagar, A.(2014). *Measurement and Evaluation*. Vinay Rakheja, Lall
     Book Depot

Cohen, L. Manion, L., & Morrison, K. (2018). *Research Methods in Education*. (8th Eds.)
     Routledge.

Davies, M. (2011). Concept mapping, mind mapping and argument mapping: what are the
     differences and do they matter? *High Education*, 62(3), 279-301

Ebel, R.L., & Frisbie, D.A. (1991). *Essentials of Educational Measurement*. (5th Eds.) Prentice
     Hall of India, New Delhi.

Kachhap, S. (2020). *A Study on the Effect of Concept mapping on Academic Achievement in
     Science for Children with Visual Impaired at Upper primary level*. Unpublished thesis,
     Faculty of Education, Banaras Hindu University.

Kalpan, R.M. & Dennis, P.S. (2002).*Psychological Testing: Principles, Application and Issues*.
     Thompson Asia Ltd, Singapore.

Miller, M.D, Linn, R.L, & Gronlund, N.E. (2009). *Measurement and Assessment in teaching*,
     (10th Eds.), Prentice Hall Inc., New Jersey: USA.

McCowan & McCowan, (1999). Item Analysis for Criterion Reference tests. Centre for
     Development of Human services, Newyork.

NCERT (2022). *Science Syllabus for IX-X*.
https://cbseacademic.nic.in/web_material/CurriculumMain22/termwise/Secondary/Science_Sec_2021-22.pdf

NCERT. (2005). *Syllabus for Secondary and Higher Secondary level classes*.
     https://ncert.nic.in/syllabus.php

Reena, R & Anisha,(2017). Construction and Standardization of Mathematics Achievement Test
     for IXth Grade Students. *Educational Quest: An International Journal. of Education
     and Applied Social Science*, 8(2), 651-655.
     https://doi: 10.5958/2230-7311.2017.00115.5

Sharma, H & Sarita. (2018). Construction and standardization of an achievement test in Science,
     International Journal of Research and Analytical Reviews, 5(3), 1037-1043.
     http://ijrar.com/

Stodola, Q., & Stordahl, K. (1972). *Basic Educational test and Measurement*, (1st Eds.) Thomson
     Press (India) Limited, Science Research Associates. Inc., USA.