

Development And Validation Of A Critical Thinking Ability Test In Science

¹Dr. P.B. Kavyakishore

¹ Assistant Professor

¹ R.V. Teachers College (IASE), Research Centre in Education

¹ Bengaluru City University, Bangalore-560 011, Karnataka, India.

Abstract: Critical thinking is an important skill in science education, enabling students to analyze information, evaluate evidence and make logical decisions. The present research aimed to construct and standardize a Critical Thinking Ability Test in Science (CTATS) for Class IX standard school children. The test was designed to measure students' ability to analyze scientific situations, apply scientific concepts and evaluate information logically. The tool consists of 25 items divided into three sections: multiple-choice questions, short-answer reasoning questions and identification of relevant information. The preliminary version of the test was administered to 100 Class IX students from 3 schools that is government, aided and private school situated in Bangalore South, Karnataka. Item analysis was conducted through difficulty index and discrimination index to determine the quality of the items. The results indicated that the items had acceptable levels of difficulty and discrimination. Reliability of the test was established using split-half reliability (0.855) and test-retest reliability (0.805) methods. Content validity was ensured through expert judgment from science educators and experienced teachers. Norms were established through mean and standard deviation to classify students' critical thinking ability levels and the findings indicate that the Critical Thinking Ability Test in Science (CTATS) is a reliable and valid instrument for assessing critical thinking ability among Class IX students in science.

Index Terms - Construction, Standardization, Validation, Critical Thinking Ability, Science, School Children.

1. INTRODUCTION

Critical thinking has emerged as one of the most important competencies required for significant learning in the twenty-first century. In the context of science education, critical thinking refers to the ability of learners to analyze scientific information, evaluate evidence, interpret data and make logical conclusions based on empirical reasoning. Science education is not limited to the acquisition of factual knowledge; rather, it involves developing inquiry skills, reasoning abilities and problem-solving capacities among learners. Therefore, the development of critical thinking skills is considered important for students to understand scientific phenomena and apply scientific knowledge in real-life contexts.

Researchers have provided several definitions and conceptualizations of critical thinking. Ennis (2015) defined critical thinking as reasonable and reflective thinking focused on deciding what to believe or do. According to Facione (2011), critical thinking is a purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation and inference. These definitions emphasize that critical thinking is a cognitive process that involves multiple higher-order mental operations such as reasoning, evaluating arguments, identifying assumptions and drawing logical conclusion. In science education, these abilities help students to interpret experimental results, assess scientific claims and make informed decisions based on evidence.

From a psychological perspective, the theoretical basis of critical thinking is closely related to Bloom's Taxonomy of Cognitive objectives. Bloom and his colleagues (1956) proposed a hierarchy of cognitive skills consisting of knowledge, comprehension, application, analysis, synthesis and evaluation. The higher levels of this taxonomy - analysis, synthesis and evaluation-represent complex

cognitive processes that form the basis of critical thinking. Later revisions of Bloom's taxonomy by Anderson and Krathwoh (2001) further emphasized higher-order thinking skills such as analyzing, evaluating and creating. These theoretical frameworks provide a strong foundation for constructing instruments designed to measure critical thinking abilities. Another important theoretical perspective supporting the development of critical thinking is constructivist learning theory. According to constructivist theorists such as Piaget (1972) and Vygotsky (1978), learners actively construct knowledge through interaction with their environment. Learning occurs when students engage in inquiry, discussion and problem solving. In science classrooms, students learn effectively when they investigate problems, interpret data and justify conclusions. These processes naturally require critical thinking skills. Thus, assessment tools designed to measure scientific reasoning must incorporate tasks that require analysis, interpretation and logical evaluation of information.

In science education specifically, critical thinking is closely linked with scientific reasoning and inquiry skills. According to Fisher and Scriven (2016), critical thinking in science involves analyzing experimental evidence, evaluating scientific explanations and interpreting data logically. Similarly, Paul and Elder (2014) highlighted that critical thinking involves clarity, accuracy, relevance, logic and fairness in reasoning. When applied to science learning, these elements encourage students to question assumptions, examine evidence critically and draw scientifically valid conclusions. In the Indian educational context, the importance of critical thinking has been emphasized in several national educational policies and curriculum frameworks and the National Curriculum Framework (NCF, 2005) highlights that science education should promote inquiry, reasoning and problem-solving abilities among students rather than mere memorization of facts.

Despite the growing emphasis on critical thinking in education, assessment practices in many schools still focus primarily on factual recall and routine problem solving. Many existing instruments measure general critical thinking ability but do not specifically address critical thinking within scientific contexts. Instruments such as the Watson–Glaser Critical Thinking Appraisal and the California Critical Thinking Skills Test are widely used to assess general reasoning abilities; however, they are not designed specifically for evaluating scientific reasoning among school students. Consequently, there is a need to develop subject-specific tools that measure critical thinking ability within the framework of science learning. Recognizing this gap, the present research aimed to construct a Critical Thinking Ability Test in Science (CTATS) specifically designed for Class IX students. The test was conceptualized to measure students' ability to analyze scientific situations, interpret information, evaluate evidence and distinguish between relevant information in scientific contexts. The development of this test was guided by theoretical frameworks of critical thinking, scientific inquiry and higher-order cognitive processes.

During the initial stage of test construction, the researcher carefully analyzed the science curriculum prescribed for Class IX students under the Karnataka State syllabus. Topics from physics, chemistry and biology were examined to identify contexts suitable for assessing critical thinking abilities. Based on this analysis, a pool of items was generated that required students to interpret scientific situations, analyze cause-effect relationships and evaluate logical arguments related to scientific phenomena. The items were designed to assess different dimensions of critical thinking such as analysis, reasoning, evaluation, interpretation and identification of relevant information. Multiple item formats were included to capture different aspects of critical thinking ability. These included multiple-choice questions requiring logical reasoning, short-answer questions assessing explanation and justification and tasks requiring identification of relevant and irrelevant information in scientific scenarios. The use of varied item formats ensured that the test measured both objective reasoning skills

and students' ability to articulate scientific explanations. Thus, the construction of the Critical Thinking Ability Test in Science (CTATS) was grounded in established theoretical frameworks of critical thinking, cognitive psychology and science education. The test was designed to align with contemporary educational goals that emphasize inquiry-based learning, analytical thinking and scientific reasoning among school children especially for Class IX students.

2. LITERATURE CITED ON TESTS RELATED TO CRITICAL THINKING ABILITIES

Over the past several decades, researchers and educational psychologists have developed a number of instruments to measure critical thinking ability in educational settings. These tools have primarily focused on assessing general reasoning abilities such as interpretation, inference, analysis and evaluation. Early studies in the field emphasized the importance of critical thinking as a fundamental cognitive skill necessary for effective learning and decision making. According to Ennis (2015), critical thinking refers to reflective and reasonable thinking that helps individuals decide what to believe or what action to take. Ennis further emphasized that critical thinking involves analyzing arguments, evaluating evidence and making logical judgments. Instruments developed during the early phase of critical thinking research were largely designed to measure general reasoning ability rather than subject-specific thinking.

One of the most widely recognized instruments for measuring critical thinking is the Watson–Glaser Critical Thinking Appraisal developed by Watson and Glaser (2008). This test measures several dimensions of critical thinking including inference, recognition of assumptions, deduction, interpretation and evaluation of arguments. The Watson–Glaser test has been extensively used in educational and organizational settings to assess reasoning ability. Although the instrument has demonstrated high reliability and validity across various populations, researchers have noted that it primarily measures general logical reasoning rather than domain-specific thinking skills. In science education, students are required to analyze scientific data, interpret experimental evidence and evaluate scientific explanations, which may not be adequately captured by a general reasoning test. Another important instrument used to assess critical thinking is the California Critical Thinking Skills Test (CCTST) developed by Facione and Facione (1992). This test was based on the Delphi consensus definition of critical thinking and measures cognitive skills such as interpretation, analysis, evaluation, inference, explanation and self-regulation and this test has been widely used in higher education to measure critical thinking skills among college students, while the instrument provides a comprehensive framework for assessing higher-order thinking, its application has been largely limited to higher education contexts. Moreover, the content of the test is not specifically related to school-level science concepts, which limits its suitability for assessing scientific reasoning among school students.

Similarly, the Cornell Critical Thinking Test developed by Ennis, Millman and Tomko (2005) is another widely used instrument designed to measure critical thinking skills among students. The Cornell test evaluates students' ability to identify assumptions, judge credibility of sources and draw logical conclusions from evidence. Although this instrument is appropriate for school-level students, the test focuses mainly on general reasoning tasks rather than discipline-specific thinking processes. As a result, it does not directly assess students' ability to interpret scientific phenomena or evaluate scientific arguments.

Researchers in science education have increasingly recognized the importance of developing assessment tools that measure critical thinking within specific disciplinary contexts. Fisher and Scriven (2016) emphasized that critical thinking in science involves interpreting experimental results, evaluating scientific claims and identifying cause-and-effect relationships among variables. These processes require students to integrate conceptual understanding with analytical reasoning. Therefore, assessment tools designed for science education must incorporate tasks that require interpretation of

scientific data, logical reasoning and evaluation of scientific explanations. Several empirical studies have highlighted the need for subject-specific critical thinking assessments. Like, Halpern (2014) argued that critical thinking is context-dependent and that individuals often perform differently when reasoning in specific domains such as science or mathematics. Halpern emphasized that critical thinking skills are better assessed when tasks are embedded in realistic contexts related to the subject matter. Similarly, Paul and Elder (2014) emphasized that effective critical thinking requires the application of intellectual standards such as clarity, accuracy, relevance and logical consistency when evaluating information.

In the Indian context, research on critical thinking has gradually gained attention in recent years, particularly in relation to science education. Several studies have examined the relationship between critical thinking ability and academic achievement among school students. Like, studies conducted by Rajendran (2008) highlighted that Indian classrooms often emphasize memorization rather than analytical thinking, which limits the development of critical thinking skills among students and suggested that both teaching methods and assessment practices must encourage reasoning, questioning and problem solving in order to develop critical thinking abilities. Similarly, research conducted by Sharma and Elbow (2010) emphasized that science learning becomes significant when students engage in inquiry-based activities and problem solving tasks. Their study indicated that students who participated in inquiry-oriented science instruction demonstrated higher levels of critical thinking ability compared to those taught through traditional lecture methods. However, the researchers also pointed out that the lack of standardized tools for measuring scientific critical thinking remains a major challenge in educational research. Another study by Gupta and Pasrija (2016) examined critical thinking ability among secondary school students and found that many existing instruments were not aligned with school science curricula. Their findings suggested that general critical thinking tests often fail to capture students' ability to analyze scientific situations or interpret scientific evidence and the researchers recommended the development of context-specific assessment tools that integrate scientific content with reasoning tasks.

A critical analysis of the existing instruments reveals several limitations. First, most standardized critical thinking tests focus on general cognitive reasoning rather than subject-specific thinking skills. Second, many tools are designed primarily for higher education populations and may not be suitable for secondary school students. Third, several instruments rely heavily on verbal reasoning tasks that do not adequately reflect the analytical processes involved in scientific inquiry. These limitations highlight the need for specialized assessment tools that measure critical thinking ability within the framework of science learning. In the context of current educational reforms in India, the need for such tools has become even more significant. The National Curriculum Framework (2005) emphasizes the development of analytical thinking, inquiry skills and scientific reasoning among students. These policies advocate moving beyond rote memorization towards competency-based learning and problem-solving approaches in science education. However, the successful implementation of these reforms requires appropriate assessment instruments that can measure students' critical thinking abilities in scientific contexts.

Considering these gaps in existing assessment tools, the present research aimed to construct and standardize a Critical Thinking Ability Test in Science (CTATS) specifically designed for Class-IX children. The test focuses on evaluating students' ability to analyze scientific situations, interpret data, identify relevant and irrelevant information and draw logical conclusions based on scientific reasoning. By integrating scientific content with higher-order thinking tasks, the CTATS attempts to provide a more valid and reliable measure of critical thinking ability in science learning contexts.

3. OBJECTIVES OF THE STUDY

The study was conducted with the following objectives:

1. To construct a Critical Thinking Ability Test in Science (CTATS) for Class IX children.
2. To standardize the test through item analysis and statistical procedures.
3. To establish the reliability of the test using appropriate reliability methods.
4. To establish the validity of the test through expert judgment.
5. To develop norms for interpreting students' critical thinking scores.

4. CONSTRUCTION AND STANDARDIZATION OF CTATS

The construction and standardization of the Critical Thinking Ability Test in Science (CTATS) was carried out through systematic and scientific procedures commonly followed in educational test development. The purpose of the test was to measure the level of critical thinking ability of school children of Class IX in the context of science learning. The development process involved several stages including identification of dimensions, preparation of a blueprint, item writing, expert validation, pilot testing, item analysis and establishment of reliability, validity and norms.

4.1 Construction of CTATS Test

The development of the Critical Thinking Ability Test in Science followed the steps outlined below.

4.1.1 Identification of Dimensions

The first step in test construction involved reviewing science textbooks prescribed for Class IX students and examining relevant literature on critical thinking and science education. Based on the analysis of curriculum objectives and theoretical frameworks of critical thinking, four major dimensions were identified:

1. Conceptual Understanding of Scientific Processes
2. Application of Scientific Knowledge
3. Analytical and Reasoning Skills in Science
4. Problem-Solving and Decision Making in Science

These dimensions represent higher-order thinking skills required for significant understanding of science concepts and scientific reasoning.

4.1.2 Preparation of Test Blueprint

A blueprint was prepared to ensure balanced representation of different dimensions of critical thinking. The blueprint also ensured that the items covered various areas of science such as physics, chemistry and biology.

Blueprint of the Test

Dimensions	Number of Items
1. Conceptual Understanding	12
2. Application of Scientific Knowledge	5
3. Analytical and Reasoning Skills in Science and	4
4. Problem-Solving and Decision Making in Science	4
Total	25

The blueprint served as a guide for item preparation and ensured that the test measured different aspects of critical thinking ability.

4.1.3 Item Writing Procedure

Based on the blueprint, an initial pool of 28 items was prepared. The items were designed to require students to analyze situations, interpret scientific information and draw logical conclusions. The items were framed by using Class IX Science content and everyday scientific situations related to topics in light; Heat and its effects; electricity; work, energy, force & motion; physiology of plants & animals; elements, compounds & mixtures; cytology & histology; diversity in plants & animals from Karnataka State Board text book. Care was taken to ensure that the language used was simple, clear and appropriate for Class IX students and finalized 25 items. The test consisted of three parts:

- **Part-A: Multiple-Choice Questions (17 items × 1 mark):** These items required students to select the most appropriate answer based on scientific reasoning.
- **Part-B: Short-Answer Reasoning Questions (4 items × 2 marks):** These items required students to explain their reasoning and justify scientific conclusions.
- **Part-C: Identification of Relevant and Irrelevant Information (4 items × 3 marks):** These items assessed students' ability to distinguish useful scientific information from irrelevant statements.

4.1.4 Experts Validation

The preliminary items were reviewed by a panel of experts to ensure scientific accuracy and relevance. The panel consisted of:

- Four experts in science education
- Three experienced secondary school science teachers working in Bangalore South.

The experts evaluated each item based on the following criteria:

- Scientific correctness
- Clarity of language
- Relevance to critical thinking
- Suitability for Class IX students

Based on their suggestions, ambiguous items were modified and redundant items were removed. After revision, 25 items were retained for the preliminary version of the test.

4.2 Standardization of CTATS Test

4.2.1 Method of the Study

The study followed a descriptive research method involving the construction and standardization of an educational test.

4.2.2 Sample

The preliminary version of the test was administered to 100 Class IX students from 3 schools of government, aided and private schools located in Bangalore South, Karnataka. The sample represented a mixed population of students to ensure broader applicability of the test. Students were given 60 minutes to complete the test.

4.2.3 Item Analysis

Item analysis was conducted to evaluate the quality of the test items. Two important indices were calculated that is one is Difficulty Index (P-value) and another one is Discrimination Index (D-value)

4.2.3.1 Difficulty Index

The difficulty index indicates the proportion of students who answered the item correctly. The formula used was:

$$P = \frac{R}{N}$$

Where:

P = Difficulty index

R = Number of students answering correctly

N = Total number of students

Items with difficulty values between 0.30 and 0.70 were considered acceptable.

4.2.3.2 Discrimination Index

The discrimination index measures the ability of an item to differentiate between high-achieving and low-achieving students. The formula used was:

$$D = \frac{(U - L)}{N/2}$$

Where:

U = Number of students in the upper group answering correctly

L = Number of students in the lower group answering correctly

Items with discrimination values above 0.30 were retained.

Table-1: Item Analysis on CTATS Test

Item No	Corrected Answers	Upper Group	Lower Group	Difficulty Index	Discrimination Index	Decision
1	68	21	12	0.68	0.33	Retain
2	64	26	11	0.64	0.56	Retain
3	61	22	12	0.61	0.37	Retain
4	61	21	11	0.61	0.37	Retain
5	65	18	9	0.65	0.33	Retain
6	62	20	11	0.62	0.33	Retain
7	62	22	12	0.62	0.37	Retain
8	65	21	11	0.65	0.37	Retain
9	62	23	15	0.62	0.30	Revise
10	66	22	12	0.66	0.37	Retain
11	61	21	8	0.61	0.48	Retain
12	62	19	11	0.62	0.30	Revise
13	66	22	11	0.66	0.41	Retain
14	64	21	12	0.64	0.33	Retain
15	60	19	11	0.60	0.30	Revise
16	52	22	9	0.52	0.48	Retain
17	61	23	10	0.61	0.48	Retain
18	66	21	11	0.66	0.37	Retain
19	64	23	12	0.64	0.41	Retain
20	60	18	6	0.60	0.44	Retain
21	51	23	13	0.51	0.37	Retain

Item No	Corrected Answers	Upper Group	Lower Group	Difficulty Index	Discrimination Index	Decision
22	64	25	12	0.64	0.48	Retain
23	63	22	12	0.63	0.37	Retain
24	65	21	11	0.65	0.37	Retain
25	68	23	12	0.68	0.41	Retain

Based on the item analysis results, all test items demonstrated acceptable difficulty levels, as their difficulty indices ranged between 0.30 and 0.70. Furthermore, the majority of the items showed good discrimination power, with discrimination indices above 0.30. However, three items had borderline discrimination values ($D=0.30$) and were therefore revised to improve their effectiveness. No items were rejected. After revision, the final test consisted of 25 items.

4.2.4 Establishing Reliability

Reliability refers to the consistency and stability of test scores. Two methods were used to establish reliability.

Method	Procedure	Technique Used	Value
Split-Half	Odd-Even Item Division	Spearman-Brown Formula	0.747; Reliability = 0.855.
Test-Retest	Re-administered after 15 days	Pearson Correlation	Reliability = 0.805

4.2.4.1 Split-Half Reliability

In this method, the test items were divided into two equal halves consisting of odd and even items. The correlation between the scores of the two halves was calculated using Spearman–Brown Prophecy Formula. (Equal and Unequal length r values are 0.747 and 0.747 respectively). The reliability coefficient obtained was 0.855, indicating high internal consistency.

4.2.4.2 Test–Retest Reliability

To determine the stability of the test over time, the same test was administered again to the same group of students after an interval of 15 days. The correlation between the two sets of scores was calculated using Karl Pearson’s Product Moment Correlation and the reliability coefficient obtained was 0.805, indicating strong stability of the test scores.

4.2.5 Validity of the Test

Content validity refers to the extent to which the test adequately represents the domain of knowledge and skills it intends to measure. To establish content validity, the items were evaluated by:

- Four experts in science education
- Three experienced secondary school science teachers

The experts examined the items for:

- Scientific accuracy

- Relevance to critical thinking
- Clarity of language
- Suitability for Class IX students

Based on their suggestions, minor revisions were made. The experts confirmed that the test adequately represented the construct of critical thinking ability in science.

4.2.6 Establishing Norms

Norms were developed using the mean and standard deviation of the scores obtained from the sample. Based on the distribution of scores, five categories were created to classify students' levels of critical thinking ability.

Norms Table for CTATS Test

CTA Category	Z scores	Score Range
Very Poor	Below – 1.51	< 14
Poor	– 1.50 to – 0.51	15 – 20
Average	– 0.50 to 0.50	21 – 26
Good	0.51 to 1.50	27 – 32
Excellent	1.51 and Above	33 >

These norms enable teachers and researchers to interpret students' scores and identify different levels of critical thinking ability in science.

5. CONCLUSION

The Critical Thinking Ability Test in Science (CTATS) was systematically constructed and standardized especially for Class IX students. The test demonstrated satisfactory levels of reliability and validity and covered multiple dimensions of critical thinking in science. This test can be effectively used by teachers and educational psychologists to assess students' critical thinking skills in science learning and it may also help to teachers identify students' strengths and weaknesses in scientific reasoning and design instructional strategies to improve critical thinking ability among students.

REFERENCES

1. Anderson, L. W., & Krathwohl, D. R. (2001). *A Taxonomy for learning, teaching and assessing*. Longman.
2. National Council of Educational Research and Training (NCERT). (2005). *National Curriculum Framework*. New Delhi.
3. Ennis, R. H., Millman, J., & Tomko, T. (2005). *Cornell Critical Thinking Test Manual*. Pacific Grove, CA: Midwest Publications.
4. Ennis, R. H. (2015). *Critical thinking: A streamlined conception*. New York: Palgrave Macmillan.
5. Facione, P. A., & Facione, N. C. (1992). *The California Critical Thinking Skills Test*. Millbrae, CA: California Academic Press.

6. Fisher, A., & Scriven, M. (2016). *Critical thinking: Its definition and assessment*. Norwich: Edgepress.
7. Gupta, M., & Pasrija, P. (2016). Critical thinking ability among secondary school students in relation to academic achievement. *International Journal of Educational Research*.
8. Halpern, D. F. (2014). *Thought and knowledge: An introduction to critical thinking*. New York: Psychology Press.
9. Paul, R., & Elder, L. (2014). *Critical thinking: Tools for taking charge of your learning and your life*. Pearson Education.
10. Rajendran, N. (2008). Teaching and acquiring higher order thinking skills: Theory and practice. *University Science Malaysia*.
11. Sharma, S., & Elbow, G. (2010). Inquiry-based learning and development of critical thinking in science education. *Journal of Educational Research*.
12. Watson, G., & Glaser, E. M. (2008). *Watson–Glaser Critical Thinking Appraisal Manual*. Pearson Education.

