

ERD 2017
Education, Reflection, Development, Fifth Edition
**COMPARISON OF ASSESSMENT METHODS OF CRITICAL
THINKING**

Necmettin Şentürk (a)*

*Corresponding author

(a) Babes-Bolyai University, Cluj-Napoca, Romania, nsenturk1@googlemail.com

Abstract

Critical thinking has been defined over time by many philosophers, educators, researchers, etc. Although definitions contain a series of common skills, there are various details. Critical thinking is important for individuals to make logical decisions possible and can make life quality better. It is a deliberate act of analysing and evaluating elements of reasoning systematically. Having all this in mind, assessing people's critical thinking skills is difficult. Main assessment methods are multiple choice tests, multiple choice questions with justification, short essays and performance tests. The most common method among them is multiple choice tests which are easy to apply. However they have some disadvantages such as the lack of comprehensiveness and providing valid results. Hence there are alternative methods, which depend on the purpose of assessment.

This paper compares a series of assessment methods for critical thinking and gives an idea for appropriate assessment of critical thinking in project based learning. For this purpose 29 high school students' critical thinking skills are investigated in project based learning by using multiple choice tests, multiple choice questions with justification, short essays and performance tests. Advantages and disadvantages of assessment methods are compared to obtain proper method to use in project based learning.

© 2018 Published by Future Academy www.FutureAcademy.org.UK

Keywords: Critical thinking, assessment, project based learning, rubric.



1. Introduction

Critical thinking is a popular term in education and even in daily life. Actually it is not new term and many philosophers, authors and educators used this term under different names and applied and taught in different ways. In modern era, Dewey (1933) stated it was an active process of consideration of belief by analysing reasons and its consequences. He named this process as “reflective thinking”.

Ennis (1987) defined the term as “a reasonable reflective thinking that is focused on deciding what to believe or do.

Brookfield (1987) extended and clarified the definition as “a process that involves identifying and challenging assumptions, becoming aware of the importance of context in creating meaning, imagining and exploring reflective skepticisms ... a reflective dimensions, more than the cognitive activity of analyzing arguments –it is emotive as well as rational.

American Philosophical Association organized several panels to discuss definition, instruction and assessment of critical thinking through two years. 46 experts, are from different disciplines, participated and contributed to the Delphi Report. They reached to consensus definition as “be purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgment is based.” (Facione, 1990).

Definitions of critical thinking share basic features. Although reaching one common definition is difficult, experts stated its common components which are defined more or less same.

Critical thinking is important for individuals to make logical decisions possible and can make life quality better. It is a deliberate act of analysing and evaluating elements of reasoning systematically (Brookfield, 1987; Elder & Paul, 2000; Paul & Elder, 2000).

1.1 Assessment of Critical Thinking

Assessing critical thinking skills is difficult as well as defining critical thinking. Purpose and objectives of assessment depend on definition and approaches of critical thinking. Various purposes for assessing students’ critical thinking skills are enlisted as diagnosing the levels of students’ critical thinking and giving them feedback, motivating students to be better, informing teachers and schools about their success, doing research about critical thinking and providing information for further educational programs which students decide to enter (Ennis, 1993). In addition to educational field, critical thinking skills are assessed to select employees (Retrieved from <http://www.linklatersgraduates.co.uk/application-process/critical-thinking-test>).

Critical thinking is assessed by using multiple choice tests, multiple choice questions with justification, short essays or case studies and performance tests. The use of assessment method depends on test makers’ purpose and the size of test takers. These methods are given with their advantages and disadvantages below.

General content based multiple choice tests are usually used to select prospect students and employees or to do research about critical thinking. On a large scale use, these tests are useful because they save time, are cheap, their validity and reliability are checked (Ennis, 1993). Well-known general content based tests are Cornell Critical Thinking Test Level X and Level Z, Watson Glaser Critical Thinking

Appraisal, (Ennis, 1993), California Critical Thinking Skills Test, California Critical Thinking Dispositions Inventory, (retrieved from <https://www.insightassessment.com/Products>) , and International Critical Thinking Test, (retrieved from <http://www.criticalthinking.org/pages/international-critical-thinking-test/619>).

General content based tests could not assess students' critical thinking skills well when the test is applied after a course. Since critical thinking is part of the course and the test may not contain anything related with the subjects in the course. Also teachers could be unfamiliar to the underlying structure and theory behind the general content based tests, and lacks the time to learn. That`s why using these tests could be seen as ineffective (Haas and Keeley, 1998).

Critical thinking essay tests are general content based and more comprehensive than multiple choice tests. The Ennis-Weir Critical Thinking Essay Test is a common essay test. They present realistic tasks and permit students to justify their responses. However they are more expensive in time and/or money and not secure (Ennis, 1993). Also CIE offers thinking skills exams which have multiple choice questions, essays and case studies. CIE assess critical thinking by using case studies which includes a scenario and open-ended questions (Retrieved from <http://www.cie.org.uk/programmes-and-qualifications/cambridge-international-as-and-a-level-thinking-skills-9694/>).

Teachers and researchers may make their own test which is could be good in terms of comprehensiveness especially for subject specific assessment. They can mix the methods and form multiple-choice test with written justification. It can be easy to prepare and check like multiple choice and be open-minded as allowing students to defend their responses even can receive credit by defending their answers well (Ennis, 1993).

Performance assessments can depend on real life cases and may use for subject specific assessment because of reducing comprehensiveness. Although using real life cases is an advantage for learner, this assessment method has disadvantageous for teachers who must devote more time for performance process and assessing performance (Ennis, 1993).

1.2. Rubrics and Critical Thinking Assessment

Formative assessments collect students` performance data in a considered instruction and students` work that will promote students` abilities and higher level thinking. That`s why, formative assessments should be authentic and multidimensional (Peverini, 2009). Rubrics are used as formative assessment tools that guide students to do quality works and teacher to evaluate their works fairly. According to Popham (1997), essential elements of rubric are evaluating criteria, quality definitions and scoring scale. Rubrics may use for performance assessments as well as open-ended questions.

Rubrics offer benefits for assessment. Tasks and criteria are categorized. Hence rubrics guide the students though the process and help the teachers to give proper feedback (Tierney & Simon, 2004).

The Foundation and Center for Critical Thinking are educational non-profit organizations and their goals are to improve education primary to university. They study and teach critical thinking and develop assessment tools such as multiple choice tests and rubric. (Retrieved from www.criticalthinking.org/files/CriticalThinkingGrid.doc)

BIE is a nonprofit educational organization and expresses its priority to help teachers and to prepare students for successful lives. It forms and shares PBL products like books, articles and critical thinking rubrics.

Universities study on assessment tools for critical thinking. Critical thinking rubric of WSU is an exemplary one. WSU searched the implementation of rubric in undergraduate courses. The mean score for courses (M=3.3) in which the rubric was used was higher than the mean score for courses (M=2.44) in courses which the rubric was not use. (Retrieved from <http://web.uri.edu/assessment/files/WSU-Critical-Thinking-Rubric-2009.pdf>)

2. Problem Statement

Assessment of critical thinking is difficult in terms of reliability and required time. Choosing a proper assessment method is important for teachers. That`s why, four methods are compared and analyzed their advantages and disadvantages.

3. Research Questions

In this study two questions are searched. “Which methods are more reliable?” and “Which methods have more effective time allowance for teachers and students?”

4. Purpose of the Study

The purpose of the study was to find out a proper method for assessing students` critical thinking skills. Four methods, multiple choice, multiple choice with justifications, short answer and rubric for performance test are applied and compared.

Subjects: 9th and 10th grade students (n=29) studied on a chemistry project in 5 weeks and presented their products in 2 weeks. All course materials and assessment instruments are in English. Groups includes 3 members, which are one high level, one medium level and one low level students in terms of the mean of first semester exams.

5. Research Methods

Learning and improvement in thinking requires more time and happens in years (Paul & Elder, 2000). Differences in background beliefs and assumptions between students and teachers and the difference, or similarities in the instruction make disadvantageous to use general content-based test (Ennis, 1993). Therefore I used disciple specific assessment tools for a 7-week study.

Students studied in PBL environment. A rubric, is designed for PBL, was used to assess students` critical thinking skills during preparation of the projects. (Retrieved from www.bie.org/object/document/6_12_critical_thinking_rubric_ccss_aligned)

After the presentations, multiple choice test, in which chemistry related questions were selected from science reasoning tests, has two parts. First part is composed of 3 cases and 5-7 multiple choice questions for each case and second part includes multiple choice questions with justification part in order to understand students` reasons for their choices.

Open-ended questions were selected from Cambridge Thinking Skills Test. There is a chemistry related case study, is told by different sources. Students answered the questions about reliability of sources and possible outcomes.

Rubric was used to assess students` performance while they were preparing their projects. Other tools were applied after their presentations. Means and standard deviations of the test results and rubric scores are calculated and are presented in Table 1.

Table 01. Means and standard deviations of assessment tools

Assessment Tools	Mean	St Dev
Multiple Choice	59.4	24.9
Multiple Choice with Justification	49.3	25.5
Case study	46.7	23.5
Rubric Score	51.2	24.3

As seen in Table 1, mean of multiple choice test (M=63.3) is the highest and mean of open-ended (M=46.7) questions is the lowest.

Mean of multiple choice with justification is between the means of multiple choice and case study and its results are given as separated in Table 2.

Table 02. Analysis of Justification part

Students` answers	Mean of Multiple Choice	Mean of Justification
Correct	52.8	45.8
Wrong	47.2	27.8
No answer	0	26.4

Several students got points by choosing correct answers without explanation. A few students got points by explaining their choices although they circled wrong option. Mean of justification part is lower than multiple choice and close to case study. There are some answers without explanation in justification part.

6. Findings

The mean values decreased by asking explanation. The results of multiple choice test without explanation are better than others. Similar questions were asked in multiple choice with justification and the mean of the test is lower than multiple choice test.

The results of open-ended questions were similar to justification part. There were some poor answers such as `no evidence`, `quite reliable`, `source A is a good evidence` etc. However some students did not explain or not support their thoughts with sources. It is unclear without an explanation whether they understood the case and answered correctly or not.

Rubrics for performance assessment guide students and teachers by stating criteria for each part of the project. Then teachers follow them to check students` act and work.

Preparation of discipline-specific multiple choice test and multiple choice test with justification take less time than discipline-specific case study. Performance assessment takes more time not only preparation but also application. Tests can be prepared in few hours and applied in a lesson. However preparation of performance assessment needs several hours for preparation and teachers are occupied with the assessment in the lessons to assess all students in all parts of projects.

Multiple choice test and multiple choice test with justification are checked fast. Checking case study takes time to read and assess each answer. Performance assessment keeps the teacher busy in all applied lessons because teacher must observe students` acts and compare with criteria in the rubric.

As time allowance to prepare the test and to assess students` responds, multiple choice test and multiple choice with justification have advantages as taking less time.

7. Conclusion

The rubric is a helpful tool for teachers who can assess students` work by checking the criteria on it. Therefore teachers provide feedback clearly for each part. They are more reliable. On the other hand, performance assessment takes up time. It can be applied for small scale classes. Case studies, including open-ended questions are more reliable and needs time but not as long as performance assessment. Multiple choice test with justification is a mixture of open-ended questions and multiple choice questions. Preparation and checking are easy like multiple choice and asking for explanation and reliability is better like case studies. Case studies and multiple choice tests with justification can be applied to small and medium scale classes. Multiple choice tests are advantageous in terms of time allowance but not reliability. These tests are good for large scale classes.

Justification part is important because teachers could not be sure that students really knew the answer and selected the option surely or they just circled the option randomly as given in Table 2.

References

- Brookfield, S.D. (1987). *Developing critical thinkers*. San Francisco: Jossey- Bass Publishers.
- Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process*. Boston: D.C. Heath.
- Elder, L., and Paul, R. (2000). Critical thinking: Nine strategies for everyday life, Part II. *Journal of Developmental Education*, 24(2), 38–39.
- Ennis, R. H. (1987). A taxonomy of critical thinking dispositions and abilities. In J. Baron & R. Sternberg (Eds.), *Teaching thinking skills: Theory and Practice* (pp. 9-26). New York: W. H. Freeman.
- Ennis, R. H. (1993). Critical thinking assessment. *Theory into Practice*, 32 (3), 179-186.
- Facione, P.A. (1990). American Philosophical Association. *Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction*. 1990. ERIC document ED 315-423
- Haas, P.F., and Keeley S.M. (1998). Coping with faculty resistance to teaching critical thinking. *College Teaching* 46:63–67.
- Paul, R., and Elder, L. (2000). Critical thinking: Nine strategies for everyday life, Part I. *Journal of Developmental Education*, 24(1), 40–41.
- Peverini, S. (2009). The value of teacher expertise and good judgment: Recent inspiring reading about assessment. *Language Arts*, 86(5), 398-402.
- Popham, J. (1997). "What's Wrong - and What's Right - with Rubrics". *Educational Leadership* 55 (2): 72–75.

Retrieved from: <http://www.cie.org.uk/programmes-and-qualifications/cambridge-international-as-and-a-level-thinking-skills-9694/>

Retrieved from https://assessment.trinity.duke.edu/documents/WashingtonStateUniversityCriticalThinkingProjectResourceGuide_000.pdf .

Retrieved from: <https://www.insightassessment.com/Products>

Retrieved from: <http://www.criticalthinking.org/pages/international-critical-thinking-test/619>

Retrieved from: <http://web.uri.edu/assessment/files/WSU-Critical-Thinking-Rubric-2009.pdf>

Retrieved from: www.bie.org/object/document/6_12_critical_thinking_rubric_ccss_aligned

Retrieved from: www.criticalthinking.org/files/CriticalThinkingGrid.doc

Tierney, R., & Simon, M. (2004). What's still wrong with rubrics: Focusing on the consistency of performance criteria across scale levels. *Practical Assessment, Research & Evaluation*.