

Performance-Based Assessment of Students' Critical Thinking Skills: The Case of Iranian MA and BA TEFL Students

Mohammad Aliakbari¹⁺, Akram Sadeghdaghighi¹

¹Ilam University- Iran

Abstract. In recent decades, the need to teach students to think critically has been emphasized and become a major concern among educators and researchers at all educational levels. Significance of students' ability in critical thinking triggered different ways of assessing the issue, a well-known yet often neglected example of which is performance-based assessment. Observing this type of assessment, the present study sought to examine the role of the instructional practice and gender on MA and BA students' critical thinking. To this end, 48 MA and BA students from two Universities in Ilam and Karaj were enrolled in the study. The results from their writings indicated that MA students obtained significantly higher scores than BA students. Besides, males moderately outperformed their female counterparts and chi-square test results confirmed the significant superiority of males over female learners.

Keywords: critical thinking; performance-based assessment; instructional practice; gender

1. Introduction

As one of the important aims of higher education, the ability of thinking critically is the most interesting and frequently discussed issue in educational systems. In universities, there is the expectation that students get mature from being a novice to being an expert in different dimensions, i.e. as Vardi (1999) claims they take a path in which they move from acquiring knowledge to critically examining the ideas.

Despite its importance, there is no unitary definition for critical thinking (hereafter, C.TH). Facione (1990) regards critical thinking to 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. In an attempt to define the term, Cottrell (2005) defines critical thinking as a cognitive activity which means thinking in the best way and using mental processes like attention, selection, judgment, etc. It makes people more precise in the way they work and think, more accurate in relevant and irrelevant issues, better decision makers about whether something is true and effective or not.

Along with the fact that incorporation of C.TH in different aspects of life has become prevalent, its assessment gained sophisticated attention. The evaluation can be accomplished by a variety of different methods: direct observation, tests, interviews, combination of these, etc. In Wal's (1999) terms two main approaches can be taken for the assessment of critical thinking: the first approach relates to assessing critical thinking in relation to other relevant academic skills, such as writing, oral presentation, or practical problem solving. The second approach pertains to assessing critical thinking skills as a trait or individual feature of the learner, by inviting the learner to complete an assessment scale.

Related to the first proposed approach of Wal (1999) is a kind of assessment called performance-based assessment, an innovative and crucial way to gauge C.TH ability. Based upon this approach, the present research tries to examine the extent to which MA and BA students majoring TEFL in two universities of Iran are critical thinkers.

2. Literature Review

⁺Corresponding author. Tel.: + 988412223399; fax: +988412238528.
E-mail address: maliakbari@hotmail.com

Importance of C.TH in all aspects of life forced different researchers to examine the issue from different perspectives. In line with the developments in the theory and knowledge about critical thinking, educators embraced different ways to assess critical thinking ability. In so doing, King, Wood, and Mines (1990) in their study investigated whether the critical thinking scores of college and graduate students would differ by educational level and gender, using tests of critical thinking that reflect different degrees of problem structure. They found significant main effects for educational level and gender on each of the three critical thinking tests, i.e. graduate students and males scored higher than the undergraduate seniors and females.

Azar (2010) investigated the difference between academic achievement in selection and placement exam for university students who had high critical thinking dispositions and of students who had low critical thinking disposition and whether this difference varied with students' gender, class levels, and the major they studied in high school. The result of this study indicated that there was no statistically significant difference between students' critical thinking dispositions and their gender, class levels, and the major they studied in high school.

The above mentioned research used multiple choice tests of C.TH. However, as Mulhal (2010) declares "the use of standardized assessments to measure critical thinking skills and critical thinking disposition attributes has been disappointing" (p. 29). Besides, in Blattner and Frazier's terms (2002) multiple-choice tests were suspect, lacking the face validity of a performance-based assessment. As a result, performance based assessment of C.TH had been developed in 1990; a movement which was flourished to ascertain critical thinking through students' writing ability. As Hobson and Schafermeyer (1994) declared "writing allows for examination and reexamination, debate and decision making, choice and revision, as well as cognitive activities which require more higher-order thinking skills of the communicator" (p. 423). An important point in evaluating thinking through writing is scoring students' writing. In their paper, Blattner and Frazier (2002) mentioned that Facione and Facione's (1994) critical thinking scoring rubric can be applied to students' writing to evaluate their abilities to accurately interpret evidence, identify salient arguments, analyze and evaluate alternative points of view, draw warranted conclusions, justify and explain results, and regard evidence with an open mind.

As a relatively new way of assessment, performance based assessment, to our best knowledge, is a neglected arena in C.TH assessment. This issue along with the possibility of determining C.TH through writing, the scarcity of the theme in Iranian educational system, and the contradictory role of gender on critical thinking invoked the authors to determine C.TH ability of male and female MA and BA students in Ilam and Karaj Universities through their writings.

3. Statement of the Problem

As a purposeful activity, critical thinking influences all aspects of human life as well as education and there is a great deal of interest in training people to be good critical thinkers. In this regard, one of the important aims of educational systems, especially higher education is developing learners' critical thinking skill. However, as Ozmen (2008) mentions the empirical studies conducted on the assessment of critical thinking have clearly shown that most of the higher education institutions may not be effective in teaching critical thinking. Different factors like misconceptions of critical thinking, the traditional teaching and learning habits like rote learning and the reservations of general educational system may lead to this problem. In line with the above-mentioned conflict in the literature and in an attempt to respond the enquiry on the success of the current higher educational system of Iran in developing C.TH among the university students, the authors of the present study aimed at the assessment of C.TH between two groups of MA and BA students. In addition to educational level, the contradictory role of gender is another controversial issue in this regard which is discussed in the study.

4. Methodology

4.1. Participants

In conducting the current research, two groups of English MA and BA students with their age ranging between 22 to 32 were involved. The study was conducted with 48 English students, divided in to two equal

groups according to their level of education. Participants who obtained their professional education in two universities of Ilam and Karaj have been selected without any special training in C.TH skills. Among forty eight students, twenty four were males and the other half were females.

4.2. Instrument

To score students' writings this study heavily relied on the four level holistic critical thinking scoring rubrics, developed by Facione and Facione (1994), which provides an effective instrument for performance-based assessment of students' critical thinking skills. According to them important points in using the rubric are as follows: 1) understand what the rubric is intended to address: this four level rubric treats critical thinking as a set of cognitive skills in which a good critical thinker engages in analysis, interpretation, evaluation, inference, explanation, and meta-cognitive self-regulation. 2) Differentiate and focus: holistic scoring requires focus on critical thinking, content knowledge, and technical skill. In scoring for any one of the three, the scorer must attempt to focus the evaluation on that element to the exclusion of the other two. 3) Practice, coordinate, and reconcile: in a training session with other raters, sample essays which are representative of each of the four levels will be examined. Raters will be asked to evaluate and assign ratings to these samples. After comparing these preliminary ratings, collaborative analysis with the other raters and the trainer is used to achieve consistency of expectations among those who will be involved in rating the actual cases.

4.3. Rating Scheme

Answering the questions needed collecting writing corpora from students. After that, each writing obtained a score between 1-4 according to the rubric. To care for subjectivity in rating, two raters attended to score the writings and they attended a session, so there was a discussion and agreement on the ratings. If two scorers disagreed on the scores attributed to each writing, there were three ways to overcome the conflict: a) by mutual conversation b) by using a third scorer, and c) by taking the average of the two initial rating. Because in this rubric, half point scoring was inconsistent with its intent, the second way was applied if necessary. One of the authors worked as the first rater and the other two raters were among MA students majoring TEFL in Ilam university. Besides, a double-blind scoring procedure was used, i.e. the first reader's score was removed from the essay before the second reader evaluated the student's work.

4.4. Procedures of Data Collection and Analysis

In order to reach the objectives of the study, a corpus with a total number of forty eight pieces of argumentative essay was gathered. In April 2011, the writings were collected in Ilam and Karaj Universities by the researchers. Students were asked to write an argumentative essay on the following topic: "should or should not university students work during their study". Care was taken that the collected corpora were nearly equal in size, between 200-250 words. Then the writings were scored according to Facione and Facione's (1994) critical thinking rubric.

Later, descriptive statistics, including frequency and percentage analysis, were used to describe to what extent each group of student used critical thinking skills in their writing. In so doing, after the extrapolation of the critical thinking aspects underlying the two corpora, frequency and percentage were computed and then chi-square was applied to show the significance of the results. Furthermore, determining gender differences in this respect needed the same statistical processes of frequency, percentage and chi-square for two groups of males and females.

5. Result

Table 1 illustrates the results obtained from MA students' writings. As demonstrated, among total number of twenty four students, nobody gained score point 0 which is indicative of the fact that MA students were not at the zero level of C.TH. Three out of twenty four students (12.5%) gained score point 4 which shows complete familiarity of these students with C.TH skills. Nine students (37.5%) obtained score point 3 which is above the average level and shows a high but not complete familiarity of this group with the skills. Besides, 12 (50%) students received score point 2 which is representative of the fact that they were at a level below the average. Score point 2 shows a moderate lack of familiarity with C.TH skills and the fact that their

present level of C.TH is not satisfactory. Generally speaking, in this group, half of the students were above the average level and the second half were below the medium.

Table 1. Frequency and percent of C.TH scores obtained by MA students

	Score				Total
	Zero level (1)	Below the medium (2)	Above the medium (3)	Perfect (4)	
Frequency	0	12	9	3	24
Percent	.0%	50%	37.5%	12.5%	100%

To screen the status of C.TH among BA students, frequency and percentage of their scores were calculated and presented in table 2.

In contrast to MA students, 12 students (50%) obtained score point 1 which indicates that they were not able to reflect C.TH skills in their writings. In this group, nobody gained 4, i.e. no one out of twenty four students was characterized as a perfect critical thinker. Besides, nine students (37.5%) obtained score point 2 and three students (12.5%) score point 3. In sum, out of twenty four, twenty one students (87.5%) were below the average level and just three students (12.5%) were above the medium. The fact that 87.5% of the BA participants were below the average level was indicative that BA students were not good critical thinkers.

Table 2. Frequency and percent of C.TH scores obtained by BA students

	Score				Total
	Zero level (1)	Below the medium (2)	Above the medium (3)	Perfect (4)	
Frequency	12	9	3	0	24
Percent	50%	37.5%	12.5%	.0%	100%

To check for the significance of the observed difference among MA and BA students a chi-square test was run. Since the level of significance (.000) is smaller than the expected level of $p < .05$, with respect to the results presented in table 3, the positive effect of educational level on the C.TH abilities of students is evident.

Table 3. Chi-square test results obtained for MA and BA students

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	18.429 ^a	3	.000
Likelihood Ratio	24.364	3	.000
Linear-by-Linear Association	16.000	1	.000
N of Valid Cases	48		

a. 2 cells (25.0%) have expected count less than 5. The minimum expected count is 1.50.

To compare the difference between results obtained for males and females, another chi-square test was computed. Table 4, which presents the obtained results of chi-square analysis in terms of gender differences at the significant level of $p < .002$, indicates for a significant difference between C.TH ability of female and

male students. It also shows that there exists a positive effect for gender on the C.TH skills of students in favor of males.

Table 4. Chi-square test results obtained for male and female students

Chi-Square Tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	15.381 ^a	3	.002
Likelihood Ratio	21.177	3	.000
Linear-by-Linear Association	13.673	1	.000
N of Valid Cases	48		

a. 4 cells (50.0%) have expected count less than 5. The minimum expected count is 3.50.

6. Discussion and conclusion

A priority for this study was to examine the critical thinking abilities of MA and BA students and to explore if there was any significant relation between C.TH students' abilities, their gender, and educational level. To accomplish the purpose of the research, two groups of MA and BA students majoring TEFL participated in this research and by the use of Facione and Facione's (1994) rubric their essays were assigned a score from 1 to 4.

One of the aims of the study was examining the extent to which MA and BA students are able to think and write critically? With respect to the observations of this study and based on the application of the Facione and Facione's (1994) rubric for critical thinking, the answer to this question is "about the average" for MA and "relatively minimum" for BA students. In other words, the level of C.TH observed for MA students in the study were indicative that some level of proficiency in this skill is indeed present. The overall scores demonstrated that 50% of students were at a level lower than medium and the other 50% were above this level. In the case of BA students, just three out of twenty four received a score above the medium and the remaining students fell at a level below the average. Nobody gained score point 4 which shows that they were not completely acquainted with critical thinking skills. Besides, fifty percent of the students obtained score point 1 which means that they were not familiar with C.TH skills yet. In this group, nine students obtained score point 2 which shows a near to the ground level of C.TH.

In order to answer research question which intended to find out "if there is any significant difference between MA and BA students" chi-square test was run. The findings of chi-square test analysis for this research question revealed a significant difference between two groups of students. Hence, it can be claimed that there was a meaningful difference between the C.TH ability of the given groups. In this regard, positive effect of higher education can be declared.

Another question of the study focused on the role of gender on C.TH abilities of students. As evidenced by the review of the literature, conflicting results were obtained in this regard. Results of this study are weighed against those that refuse the positive role of gender on C.TH abilities. Running chi-square test indicated for statistically significant difference between the results at $p < .002$ level. So, it can be concluded that males are better critical thinkers than females and attempts to improve females seem more serious in educational system.

Generally speaking, among total number of 48 students, thirty three students (68.8%) were below the average level of C.TH and the other remaining students (31.25%) were above this level. Of forty eight students, twelve students obtained score point 1 and were at zero level of C.TH while just three students demonstrated total familiarity with C.TH skills. So, with regard to this fact, despite superiority of males and MA students, it can be claimed that the majority of students were below the average level and improvement in C.TH skills is necessary in the educational system.

7. References

- [1] Azar, A. (2010). The effect of critical thinking dispositions on students' achievement in selection and placement exam for university in Turkey. *Turkish Science Education*, 7 (1), 61-73.
- [2] Blattner, N.H., & Frazier, C. L. (2002). Developing a performance-based assessment of students' critical thinking skills. *Assessing Writing*, 8, 47-64.
- [3] Carroll, D. W. (2007). Patterns of student writing in a critical thinking course: A quantitative analysis. *Assessing Writing*, 12, 213-227.
- [4] Carroll, R.T. (2004). *Becoming a critical thinker*. New York: Cambridge University Press.
- [5] Cottrell, S. (2005). *Critical thinking- Developing effective analysis and argument*. New York: Palgrave Macmillan.
- [6] Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction*. California Academic Press.
- [7] Facione, P.A. (1992). Critical thinking: what it is and why it counts. *Inside Assessment*, 2, 1-24.
- [8] Facione, P. A., & Facione, N. C. (1994). *The holistic critical thinking scoring rubric*. Millbrae, CA: The California Academic Press.
- [9] Halpern, D. F. (1999). Teaching for Critical Thinking: Helping college students develop the skills and dispositions of a critical thinker. *New Directions for Teaching and Learning*, 80 (1), 69-73.
- [10] Halvorsen, A. (2005). Incorporating critical thinking skills development into ESL/EFL courses. *TESOL Quarterly*, 11 (3), 1-6.
- [11] Hobson, E. H., & Schafermeyer, K. W. (1994). Writing and critical thinking: Writing-to-learn in large classes. *American Journal of Pharmaceutical Education*, 58, 423-427.
- [12] King, P. M., Wood, P. K., & Mines, R. A. (1990). Critical thinking among college and graduate students. *Review of Higher Education*, 13, 167-186.
- [13] Lee, K. S. (2004). *Case study learning strategies on critical thinking of undergraduate students*. Unpublished doctoral dissertation. Texas, Austin.
- [14] Limbach, B., Waugh, W., & Duron, R. (2006). Critical thinking framework for any discipline. *International Journal of Teaching and Learning in Higher Education*, 17 (2), 160-166.
- [15] Myers, B. E., & Dyer, J. E. (2006). The influence of student learning style on critical thinking skills. *Journal of Agricultural Education*, 47(1), 43-52.
- [16] Mulhall, M. L. (2010). *Quantitative measurement of critical thinking skills in novice and experienced physical therapists*. Unpublished doctoral dissertation. Capella University, Minneapolis.
- [17] Ozmen, L. S. (2008). Current state and understanding of critical thinking in higher education. *GÜ, Gazi Eğitim Fakültesi Dergisi*, 28 (2), 109-127.
- [18] Schafersman, S. D. (1991). *An Introduction to Critical Thinking*. Retrieved October 20, 2010, from <http://freeinquiry.com/critical-thinking.html>
- [19] Vardi, I. (1999). *Developing critical writers at the undergraduate level: some insights from critical thinking pedagogy and linguistics*. Paper presented at HERDSA Annual International Conference, Melbourne.
- [20] Wal, A. V. D. (1999). *Critical thinking as a core skill: issues and discussion paper*. Paper presented at HERDSA Annual International Conference, Melbourne.
- [21] Walker, S. E. (2003). Active learning strategies to promote critical thinking. *Journal of Athletic Training*, 38(3), 263-267.