English Language Teaching (ELT)
Vol. 1, No. 2, pp. 67- 89, 2014

Towards Assessing Critical Thinking Cognitive Ability in Varied University Majors: Evidence for EFL Students' Dominance

Azam Karimi[*]
Assistant Professor, Imam Khomeini International University
Maryam Eskafi
Assistant Professor of Sociology, Islamic Azad University of Gonabad

Abstract

The study presents findings from a comparative case study carried out to assess critical thinking levels and ability in some 350 Iranian university students from a range of majors. To this end, Watson–Glaser Critical Thinking Appraisal was applied to measure their critical thinking ability. The analyses of results revealed that the participants enjoyed an average level of critical thinking ability underpinning Iranian university students' rather average appreciation of the world and the subsequent decisions they make and of their ignorance of the intellectual traits of mind to be developed to their advantage. Furthermore, males turned out to outperform females as far as this ability was involved. Meanwhile, the undergraduates as well as EFL learners manifested a significant supremacy over Associate's degree and non-EFL students especially of empirical sciences as far as critical thinking tests were concerned. Moreover, as to the most frequent reason cited regarding why the students refrained from partaking in class discussions, regarded as the central method of helping students master critical thinking and communication skills (Wolcott, 2000), "fear of speaking in public" was rated the highest. The results of the study can have good pedagogical implications for all including EFL teachers, teacher trainers, syllabus designers as well as materials and curriculum developers.

Keywords: teaching, critical thinking, inference, recognition of assumptions, deduction, interpretation, evaluation of arguments

[*]Assistant professor, Department of English Language Teaching and Translation, Qazvin International University, Iran
Received on: 06/04/2014 Accepted on: 30/10/2014
Email: *azkarimi@ hum.ikiu.ac.ir*

1. Introduction

A rather recent and formidable challenge facing the Iranian educational system in both EFL and non-EFL areas is how it is supposed to keep compatible with the ever-increasing call for academic studies, survival of an autonomous way of life, and individual decision making in a rapidly varying society which demands people who can think and reason well and make shrewd and fair-minded judgments especially at the onset of the information era (Halpern, 2006).

Despite deep expression of concern over the urgent need for developing critical thinkers, inconsiderable efforts are made on the part of the schools to challenge students to think critically about academic subjects or developing reasoning skills especially in the area of EFL, where English language and culture can act as a superior authority in dominating how language learners deal with the copious materials they confront on a daily basis. CT is considered significant in the academic fields as it enables the students to scrutinize, evaluate, explain, and reconstruct their thinking, thereby slackening the risk of taking up, performing on, or thinking with a mistaken belief. Educators have agreed upon the fact that the improvement of higher order or cognitive intellectual abilities is of utmost significance and that CT "is central to both personal success and national needs" (Paul, 2004, p. 2). However, even academic education seems to have insufficient impacts on graduates' thinking critically for instance in making sensible interpretations of written materials and developing unbiased and well-founded arguments (Butler, 2012) to the effect that no cogent instruction for CT is broadly taking place yet.

As few educational authorities maintain, the thinking impoverishment overshadowing the schools stems from the sovereignty of conventional pedagogy. Research reveals that a majority of instructors still devote a good portion of the class time to presenting the materials or posing the questions requiring a mere re-accumulation of simple scientific facts and just about %1 of the class questions provoke the students to answer thoughtfully. Meanwhile, providing the students with few opportunities, teachers often discourage mindful viewpoints made on the part of the students (Shabani, 2003).

About three decades ago, in a paper entitled "*The Kind of Schools We Need"*, Eisner (1983) pointed out the problems attributed to the modern schools remarking that in today's schooling milieu only reading and writing literacy are emphasized while being inherently worthless and void of virtues in nurturing the learners. This is while, the chief concern

is not improving the reading ability but what and how the learners should read whereupon didactic values lie in upbringing of intellectuals who are capable of performing critically. Eisner also insisted that schools must tutor thinking, especially creative and critical abilities of pupils in what they see, hear, and read so as to be able to differentiate beliefs from facts, sophism from logical reasoning and merits from imperfections especially against information encroaching from a variety of sources. Students should be imparted in schools on how to acquire, that is, they should learn to be the autonomous designers of their own education and training (Shabani, 2003).

In order for realizing the critical pedagogy objectives, it is required to address an array of crucial areas including teaching, training, teaching methodologies and curriculum materials in all majors more systematically and seriously striving to integrate CT as a would-be integral ingredient in curriculums and consequently classrooms.

2. Literature Review

2.1. *Critical thinking (CT) Concept*

Recent approaches to education and training should target at cognitive processes defined in learner autonomy terms. Schools and colleges should insist on instructing students as to *how* to think rather than *what* to think. Problem solving, reasoning, judging, and decision making are indispensible for prosperous academic and social careers (Collier, Guenther & Veerman, 2002).

Reviewing the relevant literature on CT reveals a dearth of consensus on how best it should be defined and instructed within a sensible framework. This is while, most definitions and conceptualizations include a thread of commonality where many concepts drawn on by the educational reformers are used interchangeably (Halpern, 2006).

The notion *CT* can be traced back in the modern era as far as John Dewey who proposed and analyzed its essence in "how we think" (1933), where he considers the process to embrace two stages: thinking starts with reservation, uncertainty, skepticism and intricacy followed by curiosity, enquiry, and revelation. Ennis as one of the first scholars and researchers in this field called CT "reasonable, reflective thinking that is focused on deciding what to believe or do" (Ennis, 1987, p.10).

With this in mind, a variety of studies have strived to assemble the multiple aspects of CT so as to cater for a succinct and overarching definition. Commissioned by the American Philosophical Association, Facione's 1990 work, represents a generally agreed upon definition of CT extracted from an international assembly of 46 scholars and theoreticians

who consensually underscored cognitive skills as well as the dispositional dimension of CT. By the same token, Griggs, Jackson, Marek, and Christopher (1998) presented a summary of 25 definitions of CT in terms of "…a process of evaluating evidence for certain claims, determining whether presented conclusions logically follow from the evidence, and considering alternative explanations. Critical thinkers exhibit open-mindedness, tolerance of ambiguity, and a skeptical, questioning attitude" (p. 256).

2.2. *Watson-Glaser Critical Thinking Appraisal (WGCTA)*

Watson and Glaser (2009, in Ghanizadeh & Moafian, 2011) regard CT ability as including the cognitive ability empowering individuals in performing a variety of sub-skills. These sub-skills provide us with a taxonomy of subcategories which furthermore open up the possibility to assess CT ability drawing on an appraisal test used widely in a vast range of fields including education. The major standardized tests affiliated with assessing CT ability and ability such as WGCTA are deliberately developed around a set of subscales and items, which jointly and integratively represent the CT ability, including inference, recognition of assumptions, deduction, interpretation, and evaluation of arguments. According to Watson and Glaser (1980), the mentioned sub-categories of the WGCTA are "each designed to tap a somewhat differing aspect" of CT skills (p. 1). The exercises comprise problems, statements, reasonings and interpretations of the available information in the same way as those which a normal citizen in a society might encounter in his daily life while for instance reading a newspaper or magazine articles, and partaking in discussions on various issues.

Watson and Glaser (1980) assert that CT is an amalgamation of knowledge, attitudes, and practices in individuals. As unfolded by the User's Manual, the composite embraces:

1. Attitudes of inquiry that involve an ability to recognize the existence of problems and an acceptance of the general need for evidence in support of what is asserted to be true;
2. Knowledge of the nature of valid inferences, abstractions, and generalizations in which the weight or accuracy of different kinds of evidence are logically determined;
3. Skills in employing and applying the above attitudes and knowledge (p.1).

The Watson and Glaser (2009 as cited in Ghanizadeh & Moafian, 2011) factors or, in operational terms, the tested subcategories therewith CT is identified, are still globally applied to a variety of domains and studies. To illustrate, in a research study carried out by High (1991), a number of teachers and students were scrutinized. Results revealed that videotaped participants demonstrated the kind of attributes similar to the

variables of the WGCTA that embraced backing claims with evidence, projecting, assessing, and thinking flexibly. In another study, Newman, Webb and Cochrane (1995), adhering to the transcripts of class discussion through the web, came up with the same ideas implying the factors underlying WGCTA that included more evaluations and justification statements. Moreover, a good number of studies including Faravani (2006), and Hashemi and Zabihi (2012) have also applied WGCTA for measuring CT in EFL situations.

2.3. Background of the Study

The researchers, in the first place, mused over a variety of domestic research projects, theses and articles undertaken on CT skills and dispositions in both EFL and non-EFL areas. A meta-analysis of the results revealed significant findings as follows:

a) Quite a few studies resorted to merely descriptive and analytic research designs.

b) Research populations taken into account were diverse. Physical situations and participants' attributes including age and gender, definitely conducive to the scrutinies, were not dealt with much.

c) A point not to be taken for granted in all these studies was the crucial and infrastructural role played by the teachers and instructors under scrutiny and their managerial, and instructive ability in improving CT skills in class milieu (Shabani, 1991).

d) Curriculum planning was also regarded an outstanding facet in improving CT skills, requiring substantial consideration (see e.g. Mohammadyari, 2001).

e) Background and personal features were of great cruciality in affecting the extent to which the students displayed a penchant for thinking critically (see e.g. Ghasemifar, 2004).

f) All these studies focused on a variety of effective factors like class atmosphere, administrators' and policy-makers' roles, individual interests, subjects, course type and content, etc (see e.g. Raisdana, 2004).

g) Thinking styles had no influence on CT (Shabani, 2003).

h) The studies had a lot more to do with school than academic milieus.

On the other hand, one will meet a plethora of research studies in the same domain in foreign section which could be categorized under descriptive, analytic, experimental, and survey studies where the factors influencing CT skill embrace instructional settings, problem solving methods and reasoning. The advanced methods of thinking have also been compared with the conventional approaches. All these studies welcome CT as well as factors resulting in critical thinking. A review of the cited studies uncovers the fact that the majority of foreign or domestic

studies have only investigated the impact of a treatment or variable on some special factors. This is while, few have embarked on measuring and appraising CT skills in individuals.

In a study, Adams (1999) carried out a review of 20 previous studies conducted between 1977 to 1995 on the CT ability of nursing students. Accordingly, diploma students outperformed the AD students in Miller's study whereas Notarianni (1991) found no significant difference between BA and AD students. So, Adams came up with contradictory findings as to the question under study.

Later on, Adams and his associates (1999) designed another study to measure CT ability levels applying WGCTA among a total of 203 participants whereupon no statistically significant difference was found between the raw scores on the test and the subcategories of it, namely inference, recognition of the assumptions, deduction, interpretation and evaluation of arguments between the students of varying years. But statistically significant differences were found with regard to gender and the evaluation of arguments, where females performed much better than the male students.

Coskun (2001) conducted a study to assess the CT levels of nursing students in a university in Turkey. The experimental group was found out to score higher concerning the subcategory of deduction and the total test. Interestingly, age, marital status, parents' educational level and professions showed to play no role in how the two groups performed but the socio-economic status of the participants correlated highly with their CT levels. In another study carried out by Dayioglu (2003), the results revealed a moderate mean of the CT levels among the students who registered for the 2002-2003 academic year in Hacettepe University. Regarding disparate majors, there was found to be a crucial level of CT among the science students. Hashemi and Zabihi (2012) as well as Faravani (2006) also measured CT in EFL situations. The obtained results alluded to the fact that EFL students enjoy a considerable ability in thinking critically.

On this basis, the present paper intends to survey, measure, and analyze the CT skills of a number of Iranian students drawing upon WGCTA. These skills are furthermore viewed from multiple perspectives including gender, age, and major. The study also intends to elaborate on a related theme which emerges from university students' viewpoints as to the kind of factors which result in their non-participation in class discussions associated with thinking critically. Arguably, deep discussions can be applied to both EFL and non-EFL classrooms as the central method of helping students improve CT and communication skills (Wolcott, 2000). Discussion can be considered as a noteworthy tool for improving students' reasoning skills, as it gives teachers access to their

thought processes and is a chance especially in academic courses to guide students to higher levels of thinking. It can be rightly argued that class discussions across the entire curriculum are closely affiliated with improved CT skills, especially in the domain of language teaching (Tsui, 2002). Nevertheless, opportunities through which the students can have facilitated discussions are hardly ever catered for in family or academic vicinities despite their significance, development and transferability to other subjects or dimensions of lives.

2.4. Research questions

The study was framed to investigate the following questions:

1) Are there any differences across the scores of Quchan Azad University students regarding the 5 aspects of CT skills introduced by WGCTA?

2) Are there any differences across the scores of CT skills of Quchan Azad University students with regard to major, gender, and academic degree?

3) From the students' viewpoints, what factors lead on to their non-participation in class discussions associated with thinking critically?

3. Method

Regarding the various criteria in categorizing the research methodology in social sciences domain, the present study is of an extensive, grounded, cross-sectional, descriptive, and applied design type.

3.1. Participants

The statistical population of the study incorporated all students from both sexes in Quchan Azad University enrolled in 2006-2007 courses from a variety of majors. According to the permitted information from the registrar's office, a total number of 7512 students including 3566 males (47%) and 3946 females (53%) had been signed up in the cited semester. The result obtained from the sample size formula was 370 to the effect that this number of the students was required for the purpose of the study. Eventually, the researchers collected 350 questionnaires to analyze.

The sample size was composed almost equally of both sexes aged 25 on average, with 75% below the age of 25, representing 89.4 % as AD and BA students respectively. A good portion of the sample size comprised Chemical engineering students being 38.1% and TEFL, Nursing – Midwifery respectively with 11.4 and 9.4 percentages. A large part of the sample (37.5%) included the students enrolled in 2007. Overall, the students came from Mashad (42.7%), Quchan (31.5%), and other cities (25.8%). Sixty-three percent of the respondents reported on being allowed to talk at home to a good or considerable extent, while merely 36.7% replied that they were authorized to speak in classrooms.

A stratified random sampling was carried out with the field of study acting as the main criterion. To this end, in the first place, the proportion of the students in each major to the whole statistical population (7512) was computed and multiplied by the sample size (350) hereby specifying the number of students required regarding each single major. This number was then divided between the males and females in proportion to the sex size in the relevant major. Subsequent to the provision of all the students' names, the researchers drew upon systematic randomization to determine the members of the sample. The lottery procedure was then adopted to select the first member.

3.2. Instruments

A single page was devoted to biographical profile and the self-rating of the students concerning their background which was related to the purpose of the study including their major, gender, and academic degree. The major instrument through which the data were collected consisted of a standardized questionnaire known as WGCTA which embraced some multiple-choices as well as an open-ended question which was determined to be replied to as self-reports for evaluating classroom discussions and practices. The 80-question test was rendered into Persian with few changes made in some items in order for attending to cultural assimilations. The test has been devised around a set of 5 skills or better still sub-scales whereupon the theory of CT is represented as a whole. WGCTA classifies the sub-skill items (Watson & Glaser, 1980) as follows:

1) Inference (the ability involved in making a logical judgment based on the situational evidence to distinguish the right from the wrong in the information presented);

2) Recognizing assumptions (the ability to tell the suggested assumptions from the stated assumptions);

3) Interpretation (the ability to process information as well as determining whether generalizations on the data are warranted);

4) Deduction (the ability to reason from the general to the particular based on the information given in which the subject determines if certain conclusions necessarily follow the information provided);

5) Evaluating assumptions (the ability to distinguish between the weak and irrelevant, and strong and relevant arguments in a variety of situations) (Watson & Glaser, 1994).

The internal consistency for WGCTA, reported by the researchers were .73, .81, .85, .63 and .76 respectively alluding to the reliability of the test. And as to determining the validity of the test, while the criterion-related validities reported by quite a few researchers differed a lot, Watson and Glaser spoke of Cronbach or what they pointed to as the criterion-related validity of .60 which depicts what they refer to as a

definite practical value (Watson & Glaser, 1994). As to the Persian version of this test, some rather exact equivalents were offered under the supervision of the authorities in the field (Mohammadyari, 2001). Moreover, the researchers validated the Persian test through correlating the subdivisions of the test and total CT scores with the course grades. The correlation between the CT scores and the grades was r= .57 (p< .01)

Table 1
WGCTA Taxonomy

Number	*Categories*	*Number of items*	*Questions*
1	Inference	16	1-16
2	Recognizing assumptions	16	17-32
3	Deduction	16	33-48
4	Interpretation	16	49-64
5	Evaluating assumptions	16	50-80

3.3. Data Collection Procedure

The questionnaires were distributed among the students at the end of the semester. About 30 percent of class time was spent on providing a verbal explanation of study, encouraging involvement and ensuring anonymity plus student participation. These procedures were followed to ensure a high return rate.

4. Results and Discussion

4.1. Descriptive Data on WGCTA Subtests

4.1.1. Inference

The inference index is the trait constituting the foundation of the first 16 questions. The results obtained signify that the average inference grade- namely mean- was 4.95 (SD=1.8) with scores spread on a rather large range of 1 to 10. Consequently, the obtained measure was evaluated low.

4.1.2. Recognizing Assumptions

The results point to the mean of 11.5 (SD =1.63). As the scores ranged largely over a 1 to 15 continuum, with 6 and 15 as the lowest and highest scores, it could be inferred that the obtained mean score was higher than the average and subsequently evaluated high.

4.1.3. Deduction

The results obtained indicate an average total score of 9.4 (SD =1.95) with 4 and 15 as the lowest and the highest scores to the effect that the mean score is somewhat evaluated as being high.

4.1.4. Interpretation

A mean score of 10.05 and standard deviation of 1.8 was obtained. As the scores of this subtest were supposed to range largely over a 1 to 16

continuum, the obtained mean is assessed high in value. Meanwhile, the obtained scores were between 4 and 15.

4.1.5. Evaluation of Assumptions

The results obtained based on this 16-item subscale indicated that the students obtained a total average score of 9.24 (SD =1.99) on the subtest ranging from 1 to 16. As a result, the mean score is assessed high with the low and high scores of 3 to 14.

4.1.6. The Total CT Test

The total score obtained from adding the subtests resulted in an average total score of 44.36 and the standard deviation of 4.79. As there was a rather wide range of 1 to 80 considered for the distribution of scores, overall the mean score is appraised as average with the students' scores spread over a 29 to 60 continuum.

To investigate the normality of the distribution of WGCTA scores against a normal curve, the Kolmogorov-Smirnov test, which is the chief test for determining uniform data sets, was adopted. The results signified that the CT variable enjoys a normal distribution ($p > .05$). The following chart compares the distribution of the variable Y- hereby CT- against the normal distribution.

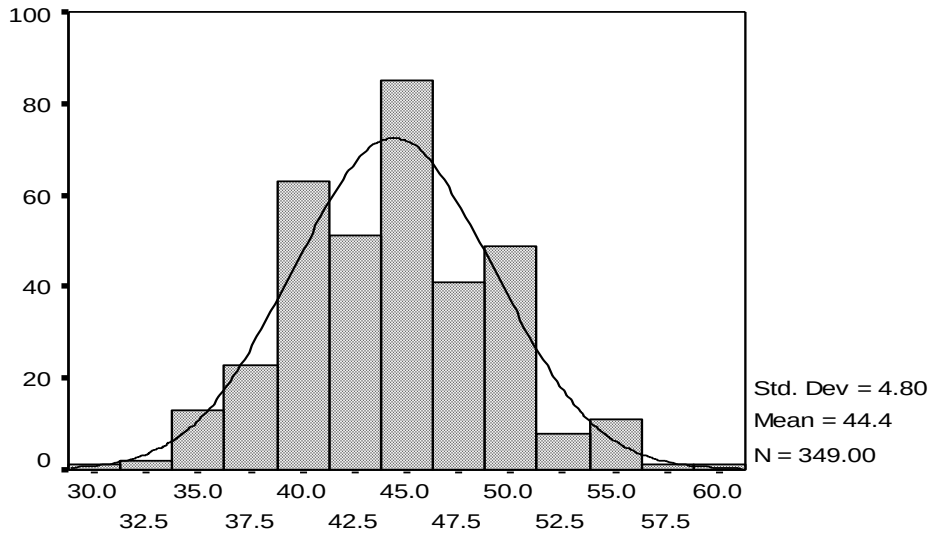

Figure 1. The normality test of CT distribution

4.2. Demographics and WGCTA Scores

In this part, the researchers compare and investigate the CT scores with regard to gender, academic degree, and major.

4.2.1. Gender

For the purpose of comparing the obtained scores of males and females, an independent samples t-test was conducted. The relevant tables are presented below:

Table 2

The Independent Samples T-Test by Gender and CT

Gender	*No.*	*Levene's Test for Equality of Variances*	*Mean*	*F*	*sig*	*T*	*Sig (2-tailed)*
M	169	Equal variances assumed	45.20	.59	.44	3.21	.001
F	181	Equal variances not assumed	43.57			3.20	.001

Table 3

The Independent Samples T-Test by Gender and CT Subtests

CT subtests	*Gender*	*Levene's Test for Equality of Variances*	*Mean*	*F*	*sig*	*T*	*sig (2-tailed)*
1.Inference	M	Equal variances assumed	5.26	4.05	.04	3.07	.002
	F	Equal variances not assumed	4.65			3.05	.002
4.Interpretation	M	Equal variances assumed	10.31	.85	.35	2.56	.01
	F	Equal variances not assumed	9.80			2.57	.01
5.Evaluating Assumptions	M	Equal variances assumed	9.48	.03	.86	2.15	.03
	F	Equal variances not assumed	9.02			2.16	.03

The results suggest that gender has a significant relationship with CT index as it has a meaningful influence on the subtests of Inference, Interpretation and Evaluating assumptions. The effects were however rejected as to Recognizing assumptions and Deduction subscales. All in all, males scored significantly higher than females with respect to CT scores (p=.001) and the three cited components with .002, .001, and .03 levels of significance respectively.

4.2.2. Academic Degree

According to Table 4, there was found to be no effects on the part of the academic degree on the subscales of Inference, Recognizing assumptions, and Deduction. The hypotheses however on the influence of academic degree on Interpretation and Evaluating assumptions were accepted with

the .001 and .004 level of significance respectively (see Table 5). In the end, the hypothesis testing the effect of the academic degree on CT skills was confirmed to the effect that the average CT score belonging to the students of Bachelor's degree exceeded that of AD students (p= .004).

Table 4
The Independent Samples T-Test by Academic Degree and CT Subtests

Academic degree	*No.*	*Levene's Test for Equality of Variances*	*Mean*	*F*	*sig*	*T*	*Sig. (2-tailed)*
AD	38	Equal variances assumed	42.21	.05	.81	2.90	.004
Bachelor's degree	312	Equal variances not assumed	44.61			2.94	.005

Table 5
The Independent Samples T-Test by Academic Degree and CT Subtests

CT subtests	*Academic degree*	*Levene's Test for Equality of Variances*	*Mean*	*F*	*sig*	*T*	*sig (2-tailed)*
1.Interpretation	AD	Equal variances assumed	9.43	.75	.51	2.48	.001
	Bachelor's degree	Equal variances not assumed	10.87			2.59	.001
2. Evaluating assumptions	AD	Equal variances assumed	8.21	.11	.66	2.46	.004
	Bachelor's degree	Equal variances not assumed	9.82			2.13	.004

4.2.3. Major

The findings on varying majors revealed the existence of a range of significant effects of the variable Major on the subscales of the CT test. This signifies that distinct majors enjoy various degrees of CT skills. In the present study, the highest score was found out to belong to EFL students with the lowest scores related to the students of empirical sciences.

In order to learn whether there was any influence on the part of Major on CT skills of the participants, the F test (ANOVA) was conducted. According to Table 6, the mean scores of some CT sub-skills, namely Deduction and Interpretation, were significant (p>.05). To avoid the unnecessary details, just some sub-scales are presented in Table 6.

Table 6
The F- Test (ANOVA) by Major and CT Subtests

(Sub)test		*Sum of squares*	*Degree of freedom*	*Mean square*	*F*	*Sig.*
Deduction	Between-group variance	77	9	8.614	2.303	.016
	Within-group variance	1238.055	331	3.740		
	Total	1315.578	340			
Interpretation	Between-group variance	65.506	9	7.278	2.146	.025
	Within-group variance	1122.834	331	3.392		
	Total	1188.340	340			
CT	Between-group variance	397.850	9	44.206	1.947	.045
	Within-group variance	7515.059	331	22.704		
	Total	7912.909	340			

The following chart reveals in the first place that the average means for various majors are different from one another. Secondly, the EFL, and empirical sciences majors display the highest and lowest average scores respectively.

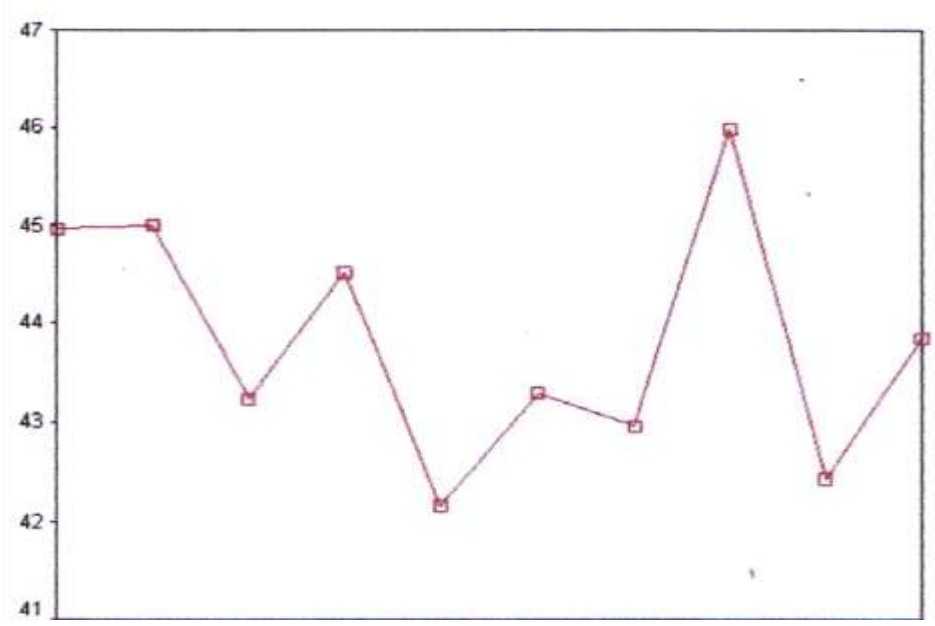

Figure 2. Different means for majors in CT terms

Note. C.E.= Chemical engineering, C.= Computer, M.A =Mechanical engineering, A.G.=Agricultural engineering, E.S. = Empirical sciences, S.S.=Social sciences, P.S.= Physical sciences, EFL=Teaching English, P.E. =Primary education, N.M.= Nursing-midwifery

In order to know which majors displayed significant differences, the LSD test was conducted. Table 7 illustrates a one-on-one comparison of the majors for merely two sub-skills of CT namely Deduction and Interpretation and evades mentioning the other subtests with no significant difference in analysis of variances. To illustrate, there was a significant difference between the variances of chemical engineering and EFL students with regard to the Deduction subtest in that EFL students outperformed the former group (p=.004).

Table 7
LSD Test for Different Majors

	LSD		*Mean difference*	*Standard error of measure-ment*	*Sig.*	*Level of confidence (95%)*	
	Majors and subtests					*Lower bound*	*Higher bound*
D	Chemical engineering	EFL	-1.03	.35	.004	-1.73	-.34
	Computer	Empirical sciences	1.61	.72	.02	.19	3.02
	Mechanical engineering	Empirical sciences	2.27	.85	.008	.60	3.95
	Mechanical engineering	Physical sciences	1.45	.73	.04	.004	2.90
	Agricultural engineering	Empirical sciences	1.40	.66	.03	.102	2.71
	Empirical sciences	Social sciences	-1.71	.694	.01	-3.07	-.34
	Empirical sciences	EFL	-2.01	.638	.002	-3.26	-.75
	Empirical sciences	Primary education	-1.16	.69	.09	-2.54	.21
	Physical sciences	EFL	1.19	.47	.01	.25	2.12
	Nursing	EFL	-1.12	.46	.01	-2.03	-.22
I	Chemical engineering	Mechanical engineering	-1.38	.63	.03	-2.63	-1.34
	Chemical engineering	Primary education	1.11	.43	.01	.26	1.96
	Computer	Mechanical engineering	1.72	.75	.02	.24	3.20
	Computer	Primary education	1.45	.59	.01	.28	2.61
	Mechanical engineering	Agricultural engineering	-1.70	.70	.01	-3.08	-.32
	Mechanical engineering	EFL	-1.73	.68	.01	-3.07	-.39
	Agricultural engineering	Primary education	1.43	.52	.00	.39	2.47
	Empirical sciences	EFL	-1.26	.60	.039	-2.45	-.06
	EFL	Primary education	1.46	.49	.004	.48	2.44
O	Chemical engineering	Physical sciences	2.01	.978	.041	.08	3.93
	Chemical engineering	Primary education	2.54	1.12	.024	.34	4.75
	Empirical sciences	EFL	-3.80	1.57	.016	-6.90	-.71
	Social sciences	EFL	-2.70	1.27	.034	-5.20	-.20
	Physical sciences	EFL	-3.00	1.16	.010	-5.30	-.71
	Empirical sciences	Primary education	3.54	1.28	.006	1.00	6.08

Note. D=deduction, I=interpretation, O= The Overall CT Test

4.3. Students' Participation in Class Discussions

An open-ended question enquired the participants to produce three self-reported reasons as to why they did not tend to take part in class discussions. A follow-up table shows the taxonomy of the results obtained:

Table 9

Justifications for Students' Non-Participation in Discussion

Reasons mentioned	*Frequency*	*Percentage*
Boredom, lethargy and fatigue	59	17
Lack of interest in major and class discussions	166	47
Talks not welcomed in educational or domestic milieu	28	8
Lack of sufficient information about the topics discussed	177	50
Feeling non-secure after discussion (especially in general courses) and after-effects	10	3
Distress and diffidence of speaking in public	232	66
Personality features (e.g. self-centeredness, pride, uncommunicativeness, disappointment, lack of cooperation spirit)	21	6
Non-understanding and lack of dynamic mind	24	7
Suffering from physical defects (e.g. stammering, talking with accents, illnesses)	8	2
Lack of positive atmosphere in classroom	116	34
Personal, domestic, mental and spiritual problems	18	5
Course type	13	4
Lack of suitable verbal communication	27	8
Professors' personal features and manners	144	41
Others (e.g. diverse attitudes, non-requirement, individual study, unpunctuality, dissatisfaction)	28	8

As Table 9 illustrates, "distress and diffidence of speaking in public" was stated as the most crucial reasoning for the respondents' non-participation in class discussions. 'Lack of sufficient information about the topics discussed', 'Lack of interest in major and class discussions', and 'Professors' personal features and manners' were pointed out as further justifications with 50, 47, and 41 percents respectively.

5. Conclusion and Pedagogical Implications

The present study aimed at identifying any differences across the scores of Quchan Azad University students with respect to the 5 aspects of CT skills introduced by WGCTA test, namely inference, recognition of assumptions, deduction, interpretation, evaluation of arguments, as well as the differences across the scores of CT skills of the same students regarding major, gender, and academic degree. The study also targeted at figuring out the kind of factors which lead on to these students' non-participation in class discussions associated with thinking critically.

According to the results obtained, of the five CT sub-skills, the mean score of Inference for all the participants was computed to be 4.9 out of 16 which is rated as *weak*. These findings are against Dayioglu's (2003) research where the lowest mean was obtained from Evaluation of arguments i.e. not good at "distinguishing between arguments which are strong and weak" (Watson & Glaser, 1980, p.2). The Iranian result by and large depicts students as having poor perception and recognition of connections and details. As they are very likely to have difficulty comprehending a text, they will consequently experience trouble inferring the relevant and irrelevant facts of it.

The mean score of Recognizing assumptions skill was 11.57 out of 16 showing a better stand compared with other sub- skills. This finding was against Dayioglu's (2003) and Coskun's (2001) results where the subjects got the higher scores from interpretation. In our case, the average score obtained by the students on Interpretation was 10.05 out of 16, meaning that the students enjoyed an average ability to process and validate information. This might allude to the fact that Iranian students are in an average position regarding "weighing evidence and distinguishing between generalizations from given data and generalizations to be warranted beyond a reasonable doubt" (Coskun, 2001, p.55). Regarding the mean score of Deduction (9.04 out of 16), where the learners should be able to come up with some general conclusions on the basis of the available information, the students seemed to be in an average and at the same time better position compared with Inference. Overall, with respect to the scores on Inference, Deduction and Evaluating assumptions which seem to be at a very low stance compared with Interpretation and Recognizing assumptions, the average scores obtained for the latter which are regarded as more complicated skills are deemed normal. These findings can provide the teachers with some sensible insights into how best they can instruct and be understood by their students and how this ability may be ameliorated through repeated drills performed in classroom. All in all, the total mean of the scores on CT was 44.3 out of 80, which is evaluated as *average* pointing to the fact

that the participants under scrutiny enjoyed an average ability in CT sub-skills as a whole. These results seem to go with Dayioglu's (2003) study where the students showed a moderate mean of the CT level.

The participants' gender also demonstrated a significant relationship with CT. This is while, the average total score of CT was more on the part of males in contrast to females. This result was in contrast to Dayioglu's (2003) study where gender turned out not to impact the CT ability levels in any significant ways. Claytor also (1997) found sex to be independent from CT skills as measured by nursing studies referred to as AMNCTI. In contrast, other studies like the ones carried out by Rudd, Baker, and Hoover (2000) and Walsh (1996) have come up with a significant relationship between CT and gender where females were found out to possess higher levels of CT. According to the scant and at the same inconclusive background studies on this matter one can no way figure out the reason why these varieties of findings exist. The result of the present study might have to do with the culture-specific nature of rather high CT levels in male participants to the effect that there exists a kind of rather aggressive gender-opinionated grounding among Iranian males especially in the area under scrutiny where there is little commitment to eradication of patriarchy and recognizing females' equal social rights.

Furthermore, the research showed higher CT scores for EFL students comparing with those of empirical sciences students keeping the lowest position in this respect. This is while, Dayioglu (2003) found a crucial level of CT among the science students. Our finding was also not consistent with Kaya's (1997) study conducted. According to her, the CT scores of both engineering and health departments were comparatively higher than those of the social science and science departments. The reason Dayioglu (2003) puts forth as to the apparent supremacy of science students regarding CT skill has to do with the nature, that is, the content and the structure of WGCTA which refers mostly to analytical and mathematical analyses. The present research however revealed a higher level of CT score for EFL students under study. The reason might have to do with what Atkinson argued in his much cited article in 1997. He asserted that CT is indeed a social practice that embraces Western cultural values not suitable for non-western students. The latter are acclimatized to a different value system, that is, collectivism and holisticism, vs. individualism of the western world, which in their own right led to a repression of individuality (Nisbett, 2003; Davidson, 1998).

Also the application of the new approaches to TEFL has resulted in creating autonomous learners and thinkers who learn to think for themselves and perform on their own besides improving problem solving skills. It is generally assumed that students master these skills backed by

the student-centered constructivist model of scaffolding adopted to both teaching and learning (Crebbin, 2004), according to which students are encouraged to become active, interactive and reflective learners. It seems that EFL tasks combined with English themes offered by teachers to meet the interactive, intellectual and linguistic challenges stimulate critical thinking in EFL learners. According to Martinez and Jimenez Nino (2013), the majority of the tasks associated with CT require interaction, reflection, discussion, arguments and role plays, which seem to be the distinctive features of the Iranian EFL classes especially in academic milieus. These tasks are normally used to foster communicative competence and interactive skills.

Furthermore, some other tasks requiring the inference of information from a variety of passages oral or written, comparing and contrasting ideas, identifying the pros and cons relating to different issues, telling facts from opinions, writing papers using the personal viewpoints and supporting arguments, conducting research studies and analyzing the collected information, judging events and people's viewpoints about them, and thinking over social effects of various issues on the communities are among the tasks known to improve CT skills (Pineda, 2003). These tasks are also applied very much in EFL circumstances. They embrace some useful activities and learning experiences which students can widely apply to state viewpoints and put in practice their self-determination, action taking, and decision making (Martinez & Jimenez Nino, 2013). In contrast to some Asian cultures like those of the Japanese and Korean who are typically identified as having trouble voicing their thoughts and attitudes, Iranian students have turned out to get highly opinionated within the last three decades being highly skilled in evaluating, producing attitudes, and opinions and taking action.

Regarding the third research question which had to do with the students viewpoints as to non-participation in class discussions and based on their self-reports, one main justification, namely, 'distress and diffidence of speaking in public' was stated as the most crucial reasoning for the respondents. This lack of confidence and discomfort the students feel might be associated with their fear of having to delve into the gray areas of thinking critically which at the same time requires the need for developing the skill of ambiguity tolerance which can be regarded as a crucial and integral aspect of CT. Aside from this, 'Lack of sufficient information about the topics discussed', 'Lack of interest in major and class discussions', and 'Professors' personal features and manners' were stipulated as further arguments in this regard. The results allude to the fact that even the university professors and families of students can conspicuously influence the levels of CT skills in students. So, in order

for acquiring more satisfying learning outcomes and to fulfill both academic and external expectations, it is deemed crucial to help faculty figure out the kind of problems their students are inwardly struggling with while having their own say. It could be logically argued that the need for this awareness-raising stems from the fact that these students are supposed to enter into a life of independence after completing their pre-determined amount of university work.

Last but not least, the findings of the present research project might have been presented in a far too small scope and somehow with shortcomings in design. Further studies are required to allow the drawing of far stronger conclusions.

References

Adams, B. (1999). Nursing education for critical thinking: An integrated review. *Journal of Nursing Education, 38(3),* 111-119.

Adams, M., Stover, L., & Whitlow, J. (1999). A longitudinal evaluation of baccalaureate nursing students critical thinking abilities. *Journal of Nursing Education, 38*(3), 139-141.

Atkinson, D. (1997). A critical approach to critical thinking in TESOL. *TESOL Quarterly, 31,* 71-94.

Butler, H. A. (2012). Halpern critical thinking assessment predicts real-world outcomes of critical thinking. *Applied Cognitive Psychology*, *26*, 721-729.

Cillizza, J. E. (1970). *The construction and evaluation of a test of critical thinking ability, grades 7-8*. Boston: Boston University School of Education.

Claytor, K. L. (1997). *The development and validation of an adult medical nursing critical thinking instrument (*Unpublished doctoral dissertation). Indiana University, Bloomington.

Cochran, W. G. (1977). *Sampling techniques*. New York: Wiley.

Collier, K., Guenther, T., & Veerman, C. (2002). *Developing critical thinking skills through a variety of instructional strategies* (Unpublished master's thesis). Saint-Xavier University, Chicago, Illinois.

Coskun, S. (2001). *Hacettepe universitesi hemsirelik yuksekokulu ogrencilerinin elestirel dusunme duzeyleri* (Unpublished master's thesis). Hacettepe University, Ankara, Turkey.

Crebbin, W. (2004). *Quality teaching and learning: Challenging orthodoxies.* New York: Peter Lang.

Davidson, B. (1998). Critical thinking faces the challenge of Japan. *Inquiry, 14*(3), 41-53.

Dayioglu, S. (2003). *A descriptive study on the critical thinking levels of the students at the unit of English preparatory school at Hacettepe*

University (Unpublished master's thesis). Hacettepe University, Ankara, Turkey.

Dewey, J. (1933). *How we think*. Boston: D.C. Heath.

Eisner, E. W. (1983). The kind of schools we need. *Educational Leadership, 41*(2), 48-55.

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., Ennis, J. M., & Tomko T. N. (1985). *Cornell critical thinking tests*. Midwest Publications, Pacific Grove, CA.

Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction*. The Delphi report, California Academic Press, Fullerton, CA.

Facione, P. (2006). *Critical thinking: What it is and why it counts.* California: California Academic Press.

Faravani, F. (2006). *Portfolio and critical thinking* (Unpublished master's Thesis). Ferdowsi University of Mashhad, Iran.

Ghanizadeh, A., & Moafian, F. (2011). The relationship between Iranian EFL teachers' sense of self-efficacy and their pedagogical success in language institutes. *The Asian EFL Journal Quarterly*, *13*(2), 249-272.

Ghasemifar, N. (2004). *Thinking mechanism*. Tehran: Ghasidesara.

Griggs, R. A., Jackson S.L., Marek P., & Christopher A.N. (1998). Introductory psychology texts and supplements. *Teaching of Psychology*, *25*, 254–266.

Halpern, D. F. (2006). The nature and nurture of critical thinking. In R. Sternberg, R. Roediger, & D. F. Halpern (Eds.). *Critical thinking in psychology* (pp. 1-14). Cambridge: Cambridge University Press.

Hashemi, M. R., & Zabihi, R. (2012). Does critical thinking enhance EFL learners' receptive skills? *Journal of Language Teaching and Research*, *3*(1), 172-179.

High, M. H. (1991). Assessing the effect of high school lessons in thinking skills. *High School Journal, 75*(1), 34-39.

Kaya, H. (1997). *Critical thinking appraisal of university students* (Unpublished doctoral dissertation). İstanbul University, Turkey.

Long, J. C. (2008). Teaching critical thinking in Asian EFL contexts: Theoretical issues and practical applications. Retrieved on April 2011 from http://www.paaljapan.org/resources/proceedings/ PAAL8/ pdf/ pdf022.pdf.

Martinez, Y., & Jimenez Nino, P. (2013). Implementing tasks that stimulate critical thinking in EFL classrooms. *Cuadernos de Lingüística Hispánica, 21*, 143-158.

Modjeski, R. B., & Michael, W. B. (1983). An evaluation by a panel of psychologists of the reliability and validity of two tests of critical thinking. *Educational and Psychological Measurement, 43*(4), 1187-1197.

Mohammadyari, A. (2001). *The relationship between the critical thinking skills of educational group managers and organizational change management in Ferdosi University* (Unpublished master's thesis). Ferdosi University, Mashhad, Iran.

Moss, P. A., & Kozdiol, S. M. (1991). Investigating the validity of a locally developed critical thinking test. *Educational Measurement Issues and Practice, 10*(3), 17-22.

Newman, D. R., Webb, B., & Cochrane, C. (1995). A content analysis method to measure critical thinking in face-to-face and computer supported group learning. *Interpersonal Computing and Technology, 3*(2), 56-77.

Nisbett, R. (2003). *The geography of thought: How Asians and Westerners think differently and why.* New York, NY: The Free Press.

Notarianni, M.A. (1991). *An investigation of the critical thinking ability of associate and baccalaureate degree nursing students (*Unpublished doctoral dissertation). Widener University, Wilmington, Delaware.

Pineda, C. (2003). *Searching for improved EFL classroom environments: The role of critical thinking-related tasks.* Bogota: Departamento de Publicciones Universidad Externado de Colombia.

Raisdana, F. (2004). To exploit critical thinking factors in a classroom. *Educational Technology Development Journal, 3,* 34-51.

Rudd, R., Baker, M., & Hoover, T. (2000). Undergraduate agricultural student learning styles and critical thinking abilities: Is there a relationship? *Journal of Agricultural Education, 41*(3), 2-12.

Shabani, H. (1991). *The effect of group work problem solving method on critical thinking and educational progress of Tehran's Fourth grade students* (Unpublished doctoral dissertation). Tarbiat Modarres University, Tehran, Iran.

Shabani, H. (2003). *Advanced teaching methodology: Teaching thinking skills and approaches.* Tehran: SAMT.

Tsui, L. (2002). Fostering critical thinking through effective pedagogy: Evidence from four institutional case studies. *Journal of Higher Education, 73,* 740-763.

Walsh, C. M. (1996). *Critical thinking disposition of university students in practice disciplines (nursing, education, and business) and non-practice disciplines (English, history, and psychology): An exploratory study (*Unpublished doctoral dissertation). College Park, MD, University of Maryland.

Watson, G., & Glaser, E. M. (1980). *Watson-Glaser critical thinking appraisal*. San Antonio: Psychological Corp.

Watson, G., & Glaser, E. M. (1994). *Watson-Glaser Critical Thinking Appraisal, Form S manual. San Antonio.* TX: The Psychological Corporation.

Watson, G., & Glaser, E. M. (2009). *Watson-Glaser II critical thinking appraisal: Technical manual and user's guide*. NY: Pearson.

Wolcott, S.K. (2000). Designing assignments and classroom discussions to foster critical thinking at different levels in the curriculum. In L. Borghans, W. H. Gijselaers, R. G. Milter and J. E. Stinson (Eds.), *Educational Innovation in Economics and Business V: Business Education for the Changing Workplace* (pp. 231-251). Dordrecht, the Netherlands: Kluwer Academic Publishers.