

Promoting and Assessing EFL College Students' Critical Thinking Skills through Argumentative Essay Writing

Talha A. Sharadgah

Department of English, Community College of AlKharj
Prince Sattam bin Abdulaziz University, Saudi Arabia

Rami A. Sa'di

Department of English, Community College of AlKharj
Prince Sattam bin Abdulaziz University, Saudi Arabia

Hamdi H. Ahmad

Department of English, Community College of AlKharj
Prince Sattam bin Abdulaziz University, Saudi Arabia

Abstract

The significance of this study is heightened by the fact that critical thinking (CT) is vastly seen as a major objective of higher education and the basis for the development of learning outcomes. Thus, this quasi-experimental aims at promoting and assessing students' critical thinking skills (CTSs) through argumentative essay-writing. It also investigates the correlation between CT and essay-writing skills. The main question addressed is: what is the effectiveness of promoting CTSs through argumentative essay-writing among English major students in terms of interpretation, analysis, evaluation, inference, and explanation? An instructional material was designed and implemented in classroom teaching to enhance CT. The study was conducted on 98 English major male participants enrolled in an essay-writing course at Prince Sattam bin Abdulaziz University (PSAU), Saudi Arabia. The participants were randomly assigned to either intervention ($n=49$) or control ($n=49$) groups. Quantitative-qualitative methods were employed. Pretest and posttest were applied to both groups. The Facione and Facione (1994) CT scoring rubric was utilized for assessing CTSs. Findings revealed that CT and essay-writing skills are significantly positively correlated. Assessment of students' essays denoted that the intervention group significantly surpassed the control in the five CTSs: "interpretation, analysis, evaluation, inference, and explanation" (Facione 1990, p. 8). It can be concluded that explicitly teaching CTSs through essay-writing can be effective in the development of these skills. The study recommended that further studies be implemented in different universities and also using other CT definitions and skills, and comparisons between the findings could be made.

Keywords: argumentative essay, critical thinking skills, essay-writing skills, promoting and assessing, Saudi University students

Cite as: Sharadgah, T. A., Sa'di, R. A., & Ahmad, H. H. (2019). Promoting and Assessing EFL College Students' Critical Thinking Skills through Argumentative Essay Writing. *Arab World English Journal*, 10 (4) 133 -150. DOI: <https://dx.doi.org/10.24093/awej/vol10no4.11>

Introduction

Promoting and assessing CTSs has become an urgent need to enhance the performance of students, especially in higher education. CT is not innate; it is a cognitive skill that can be acquired and developed through certain patterns of instructional training (Daniel & Auriac, 2011). CT comprises a set of skills such as inference, interpretation, evaluation (Facione, 1990), deduction (Furedy & Furedy, 1985), induction (Ennis, 1985), and recognition of assumptions (Watson & Glaser, 1994). Thus, when students are trained to employ CTSs through a set of steps, they may become creative thinkers (Facione & Gittens, 2015).

Yet, other researchers debate that CT cannot be taught learned or transferred to new contexts. For example, Willingham (2008) asserts that CTSs are domain-specific and that they are not transferrable at any time to other fields or disciplines. Willingham explains that the same student may exhibit CTSs in one situation, or 'domain knowledge', but fail to do so in another situation that seems similar. He also disputes the existence of multi-purpose CTSs.

Although there has been a debate over whether CTSs can only be taught in a specific domain of knowledge or be transferred to new contexts, the current study tends to support the earlier studies that suggested that CT could be developed through certain instructional training. The most compelling reason for this view is that previous studies, such as the American Psychological Association's Delphi Report (Facione, 1990), suggest that CT can be measured and assessed by providing operational definitions for CT and its related skills and sub-skills. Those studies also proposed scoring rubrics designed for the assessment of CTSs, and they provided detailed explanations of how to use those rubrics to assess CT.

There have been few empirical studies carried out on the correlation between CT and writing (Liu & Stapleton, 2018). Afshar, Movassagh, and Arbabi (2017) report "a significant positive correlation between CT and writing in a second/foreign language" (p. 8). They conclude that the higher CT level the students have, "the better their argumentative writing" becomes (p. 9). In a very similar way, WN, Syahri, and Simaibang (2018) indicate "that the students with the better [CT] have the better writing ability than the poor ones" (p. 64). They add, "the more critical students are, the more creatively they develop the writing ideas" (p. 64). Additionally, Goatly (2000) argues that writing activities can be the best way to teach CT. According to Goatly the existence of some sorts of argumentative and persuasive writing tasks is the reason why writing improves CT.

There are four approaches to teaching CT (Ennis, 1989, p. 5). The first is the "general approach", which involves teaching CT separately from the subject content. The second is the "infusion approach", where the instruction of CT and the subject content are combined, and CTSs are taught explicitly. The third is the "immersion approach", which results from students' immersion in a subject, and CTSs are not explicitly presented to students. Fourth is the "mixed approach". It "consists of a combination of the general approach with either the infusion or immersion approaches" (p. 5). According to Abrami et al. (2008), an excellent way to teaching CT is the "mixed approach" as it integrates the advantages of teaching CT as a separate course within a specific subject area.

There are four widely used standardized tests for the assessment of CT. First, there is the "California Critical Thinking Skills Test (CCTST)", developed to assess CTSs of "interpretation, analysis, inference, evaluation, and explanation" (Facione, 1990, p. 8). Second, the "Cornell Critical Thinking Test (CCTT)" is a multiple-choice test that entails test takers "[using] inductive and deductive processes, identify assumptions and judge the credibility of arguments" (Plath, English, Connors, & Beveridge, 1999, p. 208). Third, "the Ennis-Weir Critical Thinking Essay Test" (Ennis & Weir, 1985) is "an open-ended test of [CT] in which test-takers are asked to generate and evaluate arguments" (Ku, 2009, p. 71). The fourth test is "the Watson Glaser Critical Thinking Appraisal (WGCTA)" (Watson & Glaser, 1994). It is a multiple-choice test intended to assess five CTSs: "inference, recognition of assumptions, deduction, interpretation, and evaluation of arguments" (Bernard et al., 2008, p. 17).

Amongst many alternatives, the researchers selected "The Delphi Report" of Facione (1990) as the fundamental criterion to measure CT. The report defines CT as "purposeful, self-regulatory judgment that results in interpretation, analysis, evaluation, inference, and explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgment is based" (p. 3). This definition is considered to be a reasonable consensus conceptualization of CT, and it is recognized as integral to the development of CT by a large Delphi study that included 46 prominent thinkers in the field (Dwyer, Hogan, & Stewart, 2014). Furthermore, the definition has earned wide acceptance, and it is still being used by the APA to assess CT (Catchings, 2015).

Facione and Facione (1994) established a "Holistic Critical Thinking Scoring Rubric" (HCTSR) that can be used for grading written work, with a set of instructions about how to use it. It assesses CT based on six key abilities derived from the Delphi Report: "interpretation, analysis, evaluation, inference, explanation, and self-regulation" (Facione 1990, p. 8). The scoring rubric, which must be used by at least two raters, evaluates texts in a spectrum that includes assessment levels of significantly weak (no variety in CTSs), unacceptable (a limited variety in CTSs), acceptable (a variety of CTSs) and strong application of CTSs (a wide variety of CTSs).

Promoting CTSs has been deemed essential for teachers at all levels (Guiller, Durndell, & Ross, 2008). Although instructors may agree on its importance, previous studies show that CT "does not seem to be widely incorporated into college curricula" (Reed & Kromrey, 2001, p. 8). According to the experience of the researchers, many students are likely to fail to write reasonable essays because they do not use CTSs when they write on academic topics. Hence, the overall objective of this investigation is to promote and assess English major students' CTSs through argumentative essay-writing and to delineate the correlation between CT and essay-writing skills. The current study investigates the following questions:

RQ1: What is the correlation between CT and essay-writing skills?

RQ2: What is the effectiveness of promoting CTSs through argumentative essay-writing among English major students in terms of "interpretation, analysis, evaluation, inference, and explanation"?

This paper derives its significance from its attempt to teach and assess CTSs through essay writing and to uncover the potential influence these skills may have on students' writing. Thus, this study is needed to add to the existing literature on the correlation between CT and essay-writing skills. Additionally, the existence of debate over CT definition, skills, assessment tools, and the possibility of teaching it creates a pressing need for more studies to be conducted in the field in order to establish a comprehensive understanding of the issue.

Method

Participants

The representative sample included 98 undergraduate male students during the first semester of 2016/2017. The Department of English at PSAU was chosen purposefully for the implementation of this study. One of the faculty members volunteered to teach the course material. The participants were enrolled in an essay-writing course, and their ages ranged between 19 and 22. They were randomized to either intervention ($n = 49$) or control ($n = 49$) group. The simple random sampling method was used to place the participants in four classes (with about 24-25 students each).

The participants are Arab university students in Saudi Arabia studying English as a foreign language. They are all in their second year of study, and their general English proficiency level is intermediate (according to the preparatory year program (PYP) scores). Students have limited opportunity to speak English outside the classroom. Entry to the Department of English requires them to complete a one-year-long PYP first.

Instrument

Instructional Material

Initially, in order to teach the intervention group, the learning material for CTSs was selected from the textbook "*Critical Thinking: A student's Introduction*" by Bassham, Nardone, Wallace, and Irwin (2010), whereas the material for writing lessons was chosen from the textbook "*Introduction to Academic Writing*", by Oshima and Hogue (2007). The intervention group received learning material that consisted of eight CT skill-based lessons. Lessons were adapted to suit the students' abilities and to promote CTSs. The control group did not receive any specific training in CT, and they did not use the modified material. Table 1 below shows a comparison between the two groups in terms of treatment:

Table 1

Treatment schedule of both groups

	Pretest	Instructor	CT Training	Redesigned Material	CT Textbook	Writing Textbook	Posttest	Assessment
Intervention Group	Took the same pretest	Had the same instructor	Yes	Yes	Yes	Used the same textbook	Took the same posttest	The same assessment procedure was used
Control Group			No	No	No			

The CT textbook was selected from among many that had been reviewed due to its coverage of the fundamentals of CT in a clear, reader-friendly language. It also provides extensive

real-world examples and presents a step-by-step approach and thus should be easy for the participants to understand. Furthermore, each unit has a theme that seems most appropriate for the course requirements and would challenge students to think critically.

The Holistic Critical Thinking Scoring Rubric

The Facione and Facione's (1994) "HCTSR" was the instrument elected for the assessment in this study owing to several considerations. First, the rubric was developed using findings from the Delphi Report of the APA, which comprised 46 experts. Second, it was selected after reviewing several rubrics used by college professors for scoring students' essays. Also, it can be applied to students' writing to evaluate their abilities in engaging in "analysis, interpretation, evaluation, inference, explanation, and self-regulation." (Landis, Swain, Friehe, & Coufal, 2007, p. 138).

The HCTSR is a four-level scale. One point is awarded for the overall demonstration of 'significantly weak' in CT. Two points represent 'unacceptable' level in CT, three points for 'acceptable' level of CT, and four points for 'strong' level of CTs. In general, scores of 3 and 4 represent demonstration of CT and scores of 1 and 2 represent little or no evidence of CT. A four-point criterion scale in reference to Facione and Facione's (1994) Rubric was utilized to measure the key skills of CT.

Critical Thinking Training

The researchers prepared CT lessons on each of the CTs and their sub-skills for the intervention group. CTs were taught in an eight-week, in-class training program featuring three sessions per week. Every session lasted for 50 minutes. The instructor taught the learning material in two phases. The first phase was the instruction phase. Two sessions lasting for a total of approximately 100 minutes were assigned to teach the lessons on CT. These lessons included direct instruction, exchange of reflective discussions, modeling, and practice conducted in groups or individually. The second phase was conducted individually during the third session of each week. Students were instructed to write essays in response to assigned prompts reflecting on their experience with the class lessons.

The teaching of CTs to the intervention group involved activities of different types and levels of complexity, starting from simple CT activities and gradually progressing to more advanced ones. Once students had completed a number of training activities, they moved on to the essay-writing task. However, in the control group, no explicit CT instruction was incorporated into the regular curriculum.

When teaching the *inference skill*, for example, emphasis has been placed on gradually stepping up the level of complexity from two perspectives. The first is the stepping up of activities and their level of hardness, progressing from writing a few lines, to paragraph(s), to a weekly essay task. Second is the training of participants on the inference skill across its three steps. In the first step, *querying evidence*, students were taught how to query evidence by recognizing premises using their prior knowledge and the information presented in the task. In step two, *conjecturing alternatives*, students were taught how to go beyond the information given to formulating multiple options using the information they provided in step one. In the final step, *drawing conclusions*, the

instructor helped students deduce new conclusions from the available information/evidence and multiple alternatives. Table 2 shows a sample of the training activities.

Table 2
A sample of the inference skill training

		Step 2 <i>Conjecturing Alternatives</i>	Step 3 <i>Drawing Conclusions</i>
		Students formulate multiple alternatives.	Educe new conclusions or consequences from the available evidence.
		Students formulated alternatives such as:	Students generated conclusions such as:
	Step 1 <i>Querying Evidence</i> Students work in groups of four to recognize premises and formulate a strategy for gathering information about the images.	<ul style="list-style-type: none"> The organized neighborhood has a mayor who boasts about his neighborhood being organized. The other neighborhood has a mayor who is indifferent and does not care how his place looks. The residents of the organized neighborhood care about their neighborhood. The residents of the other neighborhood do not care about the appearance of their neighborhood. 	<ul style="list-style-type: none"> The mayor of the unorganized and dirty neighborhood has to be dismissed from his post. Residents of the unorganized and dirty neighborhood should be educated about contributing to the cleanliness of their neighborhood. The problem might be imputable to the lack of financial support, so support must be increased for the unorganized neighborhood.
Example 1: Two images present contrasting appearances of two neighborhoods.	Evidence: Students recognized: <ul style="list-style-type: none"> The two neighborhoods belong to different areas/cities. One of which is very organized and clean and the second is random and unclean. 		
Example 2: Parents wrote a letter to the editor of a local newspaper	Evidence: Students recognized: <ul style="list-style-type: none"> Their daughter was hit by a car. The driver was drunk. Their daughter was in hospital for a month. The punishment for the driver was only a \$550 fine. 	Students formulated alternatives about the fair punishment such as: <ul style="list-style-type: none"> Go to jail. License revoked. Pay additional fine. Jail & Fine. 	<ul style="list-style-type: none"> There should be stricter penalties for drivers under the influence of alcohol so that such unfortunate accidents do not happen again. The government must emphasize that drunk-driving is a crime that will not be tolerated.
Essay Writing: Once students have completed a number of inference activities, like those listed here, they moved on to the essay writing task.			

Validity and Reliability

The validity of the instrument was achieved through a group of six specialists in the field who gave suggestions and critical comments for the improvement of the instrument. Modifications inspired by the experts were made accordingly. To examine the usefulness of the content and procedures, to achieve inter-rater reliability, and to determine if any changes were necessary, a pilot study was performed during September 2016 with 32 students that were not selected for the actual study. The pilot study used the same material as the main study and had the same instructor. A test-retest method was administered at two different times to the same sample, with a two-week interval, to demonstrate the reliability of the pretest/posttest over time. The pilot study exhibited that the instrument was effective and no further modifications were required. The Pearson Correlation between the two applications was (0.89). The reliability of the instrument was supported by calculating the value of Cronbach Alpha (0.83). These values give an indication of acceptable reliability for the application of the study.

To establish inter-rater reliability, the three raters – the researchers and the instructor – held a training session with practice discussions. Every rater coded twenty essays of the pilot study, and results were compared. Inter-rater reliability was determined using Cohen's Kappa, and was computed at 0.71, 0.75, and 0.81. The scores of the main study were also analyzed to determine the reliability across raters. Kappa value for inter-rater reliability of the scores was 0.83 (raters: 1&2), 0.79 (raters: 1&3), and 0.81 (raters: 2&3). Therefore, the inter-rater reliability obtained was sufficiently high.

Procedure

Prior to the initiation of treatment to the intervention group, data were gathered from all participants in the two groups by subjecting them to the same pretest. The pretest was administered to participants to assess their level of CT and to ensure equivalence between groups. Participants were required to write an argumentative essay of five paragraphs, between 200-250 words, in which they debated the impact of the Internet on society. Subsequently, a meeting of the intervention group with their instructor was held to explain things and to answer any questions related to the material, procedure, or responsibilities. After that, the instructor started administering the lessons as planned. The students were required to submit one essay-writing task every week. Each essay-writing task had three main sequential steps: preparation, drafting, and revising.

The prewriting step: students were given a topic so that they could search for ideas about it outside the classroom time and then get ready for the next meeting. In the first session of each week, students were taught writing and CT skill-based lessons. They were given thirty minutes to study in class the lesson specified for the task of writing. Then they were given twenty minutes to exchange discussions about the lesson.

The drafting step: students would study a model essay on the topic and produce a parallel essay. They were then required to use the list of ideas they had produced earlier in the prewriting step to compose an essay utilizing the CTSs they had learned in their training lessons. This stage lasted for approximately fifty minutes.

The post-writing step: students would revise their drafts, and they would then share the drafts with peers to get feedback. Finally, students had some time to revise their drafts one more time before submitting the final essay to the instructor for evaluation.

The researchers and the instructor marked the essays and returned a copy of them to the students with comments. This formative assessment technique was used to assess the students' progress in CT and to give them feedback on their writings. More importantly, students expect feedback from their instructor (Hedgcock & Lefkowitz, 1994) every time they produce their essays. Right after completion of the training sessions, all participants were given the same posttest.

Data Analysis

The researchers and the instructor analyzed the students' essays based on the "HCTSR". Essays were evaluated on a scale from one indicating poor to four indicating excellent. The raters coded all the pretest, posttest, and the eight weekly essays independently. Regarding the few essays

(5 essays) where raters scored essays differently, discrepancies were resolved through subsequent discussion between the three assessors, and a final agreed-on score was assigned.

The essays developed by the participants were then assessed using the essay scoring rubric proposed in the students' textbook *"Introduction to Academic Writing"* (ibid). The scores were obtained based on five criteria: content, format, organization, punctuation and mechanics, and grammar and sentence structure. The obtained data were then computed using the SPSS.

Quantitative data were statistically analyzed to assess variations in CTSs between the two groups after the treatment period. Data were analyzed using the (SPSS, IBM version 16) to compute means (M), standard deviation (SD), Pearson correlation between variables, frequencies, and averages of scores, and one-way ANOVA to evaluate differences between groups. Values of $p < 0.05$ were considered statistically significant.

Findings

Means and standard deviations of CT and essay-writing skills of the pretest were compared to demonstrate initial equivalence between groups. Regarding CTSs, there was variation in mean scores of the intervention group ($M = 2.04$, $SD = .706$), and the control ($M = 2.02$, $SD = .721$). Also, there was variation in mean scores of the intervention group ($M = 46.06$, $SD = 16.08$), and the control ($M = 45.94$, $SD = 15.549$) relating to essay-writing skills. Accordingly, in Table 3, one-way ANOVA was performed to ascertain whether the variations were significant.

Table 3

Results of essay-writing and CTSs pretest, one-way ANOVA

Source	Essay-Writing Skills					Critical Thinking Skills				
	Sum of Squares	df	Mean Square	F	Sig.	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.367	1	.367	.001	.970	.010	1	.010	.020	.888
Within Groups	24015.633	96	250.163			48.898	96	.509		
Total	24016.000	97				48.908	97			

Table 3 indicates no statistically significant differences between the two groups relative to CTSs $F(1, 96) = .020$, $p = .888$, and essay-writing skills $F(1, 96) = .001$, $p = .970$. Therefore, the two groups were equal at the beginning of the experiment.

To address the first question, the researchers computed the Pearson Correlation between the sums of the posttest scores of the two variables: CT and essay-writing skills. The two variables were significantly positively correlated. For both variables, the correlation obtained was (.698), which is highly significant at 0.01 level (2-tailed).

For the second question, posttest means and standard deviations on CT and essay-writing skills were calculated. There was variation in the mean scores achieved by the intervention group

with respect to CTSs ($M = 3.02$, $SD = .750$), and the mean scores gained by the comparison group ($M = 2.6$, $SD = .705$). Similarly, there was variation in the mean scores of the intervention group regarding essay-writing skills ($M = 59.65$, $SD = 14.959$), and the mean scores obtained by the comparison group ($M = 52.88$, $SD = 14.919$). Accordingly, in Table 4, one-way ANOVA was used to examine if the variations were significant.

Table 4

One-way ANOVA results of essay-writing and CTSs posttest

Source	Essay-Writing Skills					Critical Thinking Skills				
	Sum of Squares	df	Mean Square	F	Sig.	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.367	1	.367	.001	.970	.010	1	.010	.020	.888
Within Groups	24015.633	96	250.163			48.898	96	.509		
Total	24016.000	97				48.908	97			

One-way ANOVA reveals statistically significant differences between pretest and posttest of CTSs $F(1, 96) = 6.271$, $p = .014$ and essay-writing skills $F(1, 96) = 5.722$, $p = .019$. This indicates that both CT and essay-writing skills appeared to be significantly improved by the implementation of the study.

The frequencies and percentages of the pretest and posttest scores in both groups were calculated in order to report the students' progress in each CTS. Percentages of scores 1 and 2 were combined for little or no evidence of CT, while 3 and 4 were combined into a single percentage that demonstrates the ability to use CTSs. Table 5 reports the results.

Table 5

Frequency and percentage of scores on the CTSs scoring rubric

Interpretation	Score	Pretest		Posttest	
		Frequency	Percentage	Frequency	Percentage
Control Group	1-2	36	73.45%	20	40.81%
	3-4	13	26.55%	29	59.18%
Intervention Group	1-2	36	73.46%	7	14.28%
	3-4	13	26.54%	42	85.71%
Analysis	1-2	38	77.55%	21	42.85%
	3-4	11	22.45%	28	57.15%
Intervention Group	1-2	35	71.43%	14	28.57%
	3-4	14	28.57%	35	71.43%
Evaluation	1-2	36	73.46%	18	36.74%
	3-4	13	26.54%	31	63.26%

Intervention Group	1-2	38	77.55%	9	18.37%
	3-4	11	22.45%	40	81.63%
Inference					
Control Group	1-2	42	85.7%	22	44.89%
	3-4	7	14.28%	27	55.11%
Intervention Group	1-2	41	83.66%	13	26.55%
	3-4	8	16.34%	36	73.45%
Explanation					
Control Group	1-2	26	53.05%	18	36.73%
	3-4	23	46.95%	31	63.27%
Intervention Group	1-2	29	59.19%	9	18.37%
	3-4	20	40.81%	40	81.63%

Table 5 shows that scores 1 and 2 indicating little or no evidence of CTSs were remarkably high in both groups on the CTSs pretest, whereas scores 3 and 4 demonstrating the ability to use CTSs were very low. This indicates that most students in both groups did not demonstrate CTSs in their essays in the pretest. As for the posttest scores, Table 5 demonstrates a significant increase in scores 3 and 4 for both groups. This indicates that most students in both groups showed evidence of progress in their ability to use CTSs between the pretest and the posttest. Participants in the intervention group displayed a greater ability to utilize CTSs than those in the control group.

Discussion

In addressing the first question, a clearly strong correlation was observed between CT and essay-writing skills. The two variables were significantly positively correlated. This outcome is compatible with the findings of previous studies that have revealed the positive impact of writing on promoting CT (Dreher, 1990; Fahim & Hashtroodi, 2012; Hanson, 2004; Hu, 2017). According to Dreher (1990), "writing enhances CT and produces writers who can write good compositions in one draft" (p. 152). Hu (2017) suggests that "writing asks for more CTSs, and in turn, writing is a proper media for teachers to develop student's CTSs".

The findings of research studies emphasize that writing argumentative essays is a powerful tool for teaching and assessing CTSs. Studies also emphasize that when students are trained to think critically, they develop more profound ideas and better understanding and as a result more developed writing. This conclusion is established by McKeachie (1994) and Dixon (1996), who stress the value of training students to exercise CT constantly during writing so that writing can be an expression of CT. Similarly, Niu, Behar-Horenstein and Garvan (2013) state that persistence in teaching these skills leads to tangible growth in CT capabilities and quality of essay-writing.

Regarding the second question, which is concerned with the role of argumentative essay-writing in fostering CTSs, all participants in both groups showed evidence of progress in CTSs between the pretest and the posttest. Participants in the intervention group, however, displayed a significantly higher ability to utilize CTSs in terms of "interpretation, analysis, evaluation, inference, and explanation" (Facione 1990, p. 8). The most likely reason for the improvement in both groups was that students had been exposed to the course content. This improvement asserts the role of writing in the development of CTSs even if students do not receive training in CTSs.

Interpretation was the first CTS that was assessed. As stated by Facione (1990), the *interpretation skill* is expounded as "to comprehend and express the significance of experiences,

situations, data, events, beliefs, rules, procedures or criteria." The *interpretation skill* contains three subsidiary skills: "*categorizing, decoding significance, and clarifying meaning*" (p. 13). Data from participants' essays showed that 85.71% of the participants in the intervention group and 59.18% in the control employed the *interpretation skill*. Both groups showed improvement, but the intervention group showed significantly greater improvement on the scale. Based on Hounsell (1984), interpretation is the uppermost skill to convey a viewpoint and present it supported by evidence. Interpretation is a skill that "must be learned by practice and cannot be acquired by memorizing someone else's interpretations" (Baur, 1960, p. 107).

When students were told to write an argumentative essay of five paragraphs about "the enormous positive influence of the Internet on society", *student A* demonstrated the *interpretation skill* as follows:

Thesis statement: *The influence of the Internet has caused many positive changes in the way we shop, communicate, and learn.*

Paragraph 1, topic sentence: *Online shopping can save time and energy.*

Paragraph 2, topic sentence: *The Internet has also brought together many families separated by miles.*

Paragraph 3, topic sentence: *Another positive attribute is that the Internet offers access to wide sources of information.*

Student A has showcased excellent indicators of CT by giving a clear and supportive thesis statement and intelligible topic sentences. He interpreted the task question by paraphrasing it and generating ideas relevant to the writing topic to clarify the meaning, as can be seen in his thesis statement. The student was also implementing *categorization* and *decoding significance* when he addressed the topic by describing its importance in three areas: "...in the way we shop, communicate, and learn."

Likewise, *student B* has utilized the *interpretation skill* by *decoding significance* of the problem without bias. He treated the argument in a logical way and described many advantages and some disadvantages for the influence of the Internet on society. For instance, he stated clearly that "*there are some drawbacks for the Internet.*" Additionally, in the excerpt below, the student attempted to exploit interpretation by *clarifying meaning*. He described some personal experiences and related the topic of writing to what he already knew.

Student B: *Unfortunately, there are some drawbacks for the Internet.*

Student B: *The Internet influences my life, and it makes my life easier. It helps me to complete my homework...*

The *analysis skill* is another CTS that saw improvement in the intervention group (71.43%) and the control (57.15%). Quitadamo and Kurtz (2007) discovered that participants who practiced writing developed their *analysis skills* significantly while the non-writing group did not. The *analysis skill* is expounded as the ability "to identify the intended relationships among statements,

questions, concepts, descriptions...to express beliefs, judgments, reasons,...or opinions" (Facione, 1990, p. 14). Analysis encompasses: "*examining ideas, detecting and analyzing arguments.*"

Examining ideas applies to the skill of identifying related statements, comparing and contrasting ideas, statements or concepts, and defining terms (Facione, 1990). It was evident that the participants had exploited the *analysis skill* by *examining ideas*. For example, *student C* made an effort to identify closely related information which was significant to set up his opinion. As shown in the excerpt below, he had examined the idea given and stated a comprehensible, consistent, and logically organized opinion. His thesis statement is clearly stated in the first paragraph followed by some indication of how the essay is to be developed.

Student C: *In my opinion, the Internet is the greatest invention of this time. It plays an essential role in every aspect of life, particularly in communication, entertainment, business and providing information.*

Other participants utilized the *analysis skill* by *detecting arguments*. According to Facione (1990), *detecting arguments* refers to giving statements or descriptions, for supporting or contesting some claim or opinion, or to recognize unstated premises of a claim. *Student D*, for example, identified the relationship between two different views. As shown in the excerpt below, he identified some unstated assumptions related to the topic, particularly, "*the Internet has some negative effects*". He discussed many advantages and some disadvantages of using the Internet, although disadvantages were not stated explicitly in the given question.

Student D: *The Internet has become an essential thing in our life, although it has some negative effects.*

Facione (2011) proposed argument mapping as an effective technique for *analyzing arguments*, which indicates one's logical thinking approach through reasons, claims, and conclusions. Throughout the course, participants were taught CT and essay-writing skills thought to be appropriate to map ideas for their essays. Many participants structured their essays sequentially in order to convince their readers of their arguments and help them clearly understand the relationships between ideas. Likewise, they displayed their ideas and content of the essay logically through transitions/signals within and between paragraphs. For example, in his thesis statement, *student E* clearly stated the order of his next paragraphs:

Student E: *The influence of the Internet has caused many positive changes in the way we shop, communicate, and learn*".

Evaluation was the third skill that was assessed. According to Facione (1990), evaluation is "to assess the credibility of statements...and...the strength of the...relationships among statements, descriptions, questions or other forms of representation" (p. 92). Evaluation comprises: "*assessing claims and assessing arguments*" (p. 92). The analysis of the essays showed increased CTSs in the intervention group (77.55%) and the control (63.26%). That is statistically significant in favor of the intervention group. This result suggests that the writing process assisted students in

improving their critical evaluation skills. According to Schwartz (1968), critical evaluation is a skill that demands constant practice and "repeated challenge."

Assessing claims involves recognizing the factors affecting the credibility, relevance, and acceptability of information or opinion, whereas assessing arguments is the ability to decide on the level of acceptance of a given argument, or "to determine whether an argument relies on false or doubtful assumptions" (Facione, 1990, p. 16). Blair and Johnson (1987) specified three standards of arguments: relevance, sufficiency, and acceptability. *Student F* made a comprehensible evaluation by *assessing* the credibility of the *claim* as well as *assessing arguments*. He wrote:

Student F: *In my opinion, I absolutely approve the claim that the Internet has positively affected society up to a point.*

The extract above shows that the student judged the claim to be true. As a result, he had developed a strong supportive argument for the claim that the Internet has an enormous positive influence on society. Additionally, In the process of *assessing arguments*, he clarified his level of acceptance by saying: "*I absolutely approve*".

Inference was the fourth *skill* to be developed and assessed. The intervention group (73.45%) significantly outperformed the control group (55.11%). *Inference* is defined as the process of seeking information required "to draw reasonable conclusions," to devise anticipation and hypotheses, "to consider relevant information and to deduce the consequences...Inference includes the sub-skills of *querying evidence*, *conjecturing alternatives*, and *drawing conclusions*" (Facione, 1990, p. 16).

Querying evidence is to recognize and judge information which requires support or needs to be addressed (Facione, 1990). In the process of writing, particularly the pre-writing stage, many sources of information exist and seem to be relevant to the issue. The students' ability to determine what information is relevant and credible reflected their use of CT *querying evidence*. According to Jones (1995), this can be demonstrated through students' ability to identify the dimensions of the issue they need to address, to set up a search plan to obtain information, to collect information that can lead to an understanding of the opinion, to judge whether this information is useful to have, and to determine the sufficiency of evidence to reach a conclusion.

Conjecturing alternatives involves the process of formulating various alternatives to resolve an issue, to assume a sequence of assumptions, and to develop a set of plans to achieve some goals (Facione, 1990). Halpern (as cited in Dwyer, Hogan, & Stewart, 2012, p. 235) posits that solving a problem involves the use of multiple problem statements to define the problem and identify possible goals, the generation and selection of alternatives, and the use of explicit criteria to judge among alternatives. *Student G*, for example, drew some implications of a different position. He assumed that the Internet might be used to lure young children and then assault them.

Student G: *Unfortunately, teenagers sometimes misuse the Internet by seducing and meeting young children for assault. This type of behavior is illegal and needs to be handled very strictly by authorities. Also, parents should monitor children's use of*

the Internet and warn them of some behaviors that might have negative consequences on them.

Drawing conclusions indicates the ability to make connections among relevant information and arrive at new conclusions (Mojica-Díaz & Sánchez-López, 2010). After students had presented evidence and made arguments throughout their essays, they drew reasonable conclusions. For example, *student H* deduced the consequences from available data, taking into account opinions, evidence, and judgments that were provided earlier. He also made a prediction about the future of the Internet.

Student H: *In conclusion, I agree with most views that the Internet has more advantages than disadvantages. I believe the Internet will play a crucial part in developing our life, and it will have more advantages in the future.*

The final skill was *explanation*. The essays from the participants in the intervention group indicated that 81.63% obtained 3 or 4 points in the *explanation skill*, vis-à-vis 63.27% in the control. The *explanation skill* is defined by Facione (1990) as "to state the results of one's reasoning; to justify that reasoning, and to present one's reasoning in the form of cogent arguments" (p. 18). Specifically, explanation comprises: "*stating results, justifying procedures, and presenting arguments.*"

Stating results is "to produce accurate statements, descriptions or representations of the results of one's reasoning activities" (Facione, 1990, p. 93). A demonstration of *stating results* can be seen in *student I's* essay. He included a situation in which he stated a result regarding personal information. In this situation, *student I* gave his opinion: "*personal information...can be risky,*" and in his own words he explained the result: "*companies can gain...*"

Student I: *In addition, much of personal information that is posted on the Internet can be risky in the wrong hands. For instance, companies can have access to your personal information and use them for their interests.*

Some students practiced the *explanation skill* by *justifying procedures*. Facione (1990) defines the skill as the ability to set forth all considerations "used in forming one's interpretations, analyses, evaluation or inferences, so that one might accurately record, evaluate, describe or justify those processes" (p. 18). *Student J*, for example, after discussing some disadvantages, *justified* the *procedure* of his thinking thus:

Student J: *This certainly doesn't mean the Internet must be abolished, otherwise some caution has to be exercised while surfing the web.*

Presenting arguments is the skill of giving reasons for agreeing to claims, or to argue and express objections on a claim (Facione, 1990). *Student K* argued against the Internet by saying:

Student K: *the Internet can represent a threat to children. Children can access almost outcast websites such as porn ones. Thus, children need to be monitored by their parents when using the Internet.*

Limitations & Future Research

Although the findings of the current study provide useful information and suggest future avenues of research, several limitations can be noted. First, this study is restricted to a sample at PSAU. Similar work could be repeated with students from different universities and the findings can be compared. Secondly, the current paper surveyed CTSs according to what is proposed in the definition of Facione (1990). It would be worthwhile to conduct further studies on this topic using other definitions in order to procure a greater understanding of the issue.

Moreover, this study is limited to the material used herein. Further research is needed to confirm the results using different materials. Also, this study involved male participants only. It was not possible to include female participants in the present study for cultural reasons. Further research will need to include female participants to ascertain if the findings in the present study also apply to females. Finally, because the intervention group might have received more attention from the instructor, there is the possibility of a halo effect at stake.

Conclusions

This study concluded that a highly significant correlation existed between CT and argumentative essay-writing skills. It appears that the more proficient the students are in CT, the better they are at writing skills and vice versa. The analysis of the participants' essays also demonstrated increased CTSs in the five skills: "interpretation, analysis, evaluation, inference, and explanation" (Facione 1990, p. 8). Hence, participants in both groups benefited from this study; however, the intervention group significantly surpassed the control group.

CT is a fundamental goal of higher education that college students must develop. Training students in a variety of CTSs can assist them in applying these skills to any situation in real life that requires reflection, analysis, evaluation, and planning. Instead of being simply passive recipients of information from the instructor, students will learn to become autonomous learners.

About the Authors:

Talha Sharadgah: I have a Ph.D. in Curricula and Teaching Methods from Amman Arab University, Jordan (2013). Currently, I am an assistant professor at Prince Sattam bin Abdulaziz University in Saudi Arabia. My areas of research interest are focused on educational assessment and technology integration in education. ORCID: <https://orcid.org/0000-0001-6338-9605>

Rami Sa'di: I have a Master's degree in linguistics (2004), University of Jordan. I have the IPA certificate in English phonetics (2006), University College London. I did Cambridge CELTA in 2005. My areas of research interest are computer-assisted TEFL, social media and TEFL, and L1-influenced pronunciation learning difficulties. ORCID: <https://orcid.org/0000-0002-6205-6144>

Hamdi Ahmad: currently completing Ph.D. in Linguistics, University of Aberdeen, UK. He holds an MSc in TESOL from Edinburgh University, UK. He is a lecturer in linguistics at Prince Sattam bin Abdulaziz University, KSA. He worked for AMIDEAST, Egypt and Saudi Aramco. He is interested in translation and Linguistics. ORCID: <https://orcid.org/0000-0002-7863-7752>

References

- Abrami, P. C., Bernard, R. M., Borokhovski, E., Wade, A., Surkes, M. A., Tamim, R., & Zhang, D. (2008). Instructional interventions affecting critical thinking skills and dispositions: A stage 1 meta-analysis. *Review of Educational Research*, 78(4), 1102-1134. <https://doi.org/10.3102/0034654308326084>
- Afshar, H.S., Movassagh, H., & Arbabi, H.R. (2017). The interrelationship among critical thinking, writing an argumentative essay in an L2 and their subskills. *Language Learning Journal*, 45(4), 419-433. <https://doi.org/10.1080/09571736.2017.1320420>
- Bassham, G., Nardone, H., Wallace, J. & Irwin, W. (2010). *Critical thinking: a student's introduction*. New York, NY: McGraw-Hill Education.
- Bernard, R. M., Zhang, D., Abrami, P. C., Sicol, F., Borokhovski, E., & Surkes, M. A. (2008). Exploring the structure of the Watson–Glaser Critical Thinking Appraisal: One scale or many subscales? *Thinking Skills and Creativity*, 3(1), 15-22. <https://doi.org/10.1016/j.tsc.2007.11.001>
- Blair, J. A., & Johnson, R. H. (1987). Argumentation as dialectical. *Argumentation*, 1(1), 41–56. <https://doi.org/10.1007/BF00127118>
- Baur, E. J. (1960). A student guide for interpreting case material. *Improving College and University Teaching*, 8(3), 104-108. <https://doi.org/10.1080/00193089.1960.10534108>
- Catchings, G. (2015). A practical coaching model for critical thinking skill and leadership development (C/CTSLED). *Management and Organizational Studies*, 2(4), 42-53. <https://doi.org/10.5430/mos.v2n4p42>
- Daniel, M., & Auriac, E. (2011). Philosophy, critical thinking and philosophy for children. *Educational Philosophy and Theory*, 43(5), 415-435. <https://doi.org/10.1111/j.1469-5812.2008.00483.x>
- Dixon, F. (1996). The use of the Hegelian dialectic as a heuristic in literature classes to improve critical thinking skills (Unpublished Doctoral Dissertation). Purdue University, West Lafayette, Indiana, USA.
- Dreher, J. (1990). Preservice early childhood teachers' attitudes toward the process approach to writing. *Early Child Development and Care*, 56(1), 49–64. <https://doi.org/10.1080/0300443900560106>
- Dwyer, C. P., Hogan, M. J., & Stewart, I. (2012). An evaluation of argument mapping as a method of enhancing critical thinking performance in e-learning environments. *Metacognition and Learning*, 7(3), 219-244. <https://doi.org/10.1007/s11409-012-9092-1>
- Dwyer, C. P., Hogan, M. J., & Stewart, I. (2014). An integrated critical thinking framework for the 21st century. *Thinking Skills and Creativity*, 12, 43-52. <https://doi.org/10.1016/j.tsc.2013.12.004>
- Ennis, R. H. (1985). A logical basis for measuring critical thinking skills. *Educational Leadership*, 43(2), 44–48.
- Ennis, R. H. (1989). Critical thinking and subject specificity: Clarification and needed research. *Educational researcher*, 18(3), 4-10.

- <https://doi.org/10.3102/0013189X018003004>
- Ennis, R. H., & Weir, E. (1985). *The Ennis–Weir Critical Thinking Essay Test*. Pacific Grove, CA: Midwest Publications.
- Facione, P. (1990). *Critical thinking: a statement of expert consensus for purposes of educational assessment and instruction* (The Delphi Report). Retrieved from ERIC database, document No. ED 315423. Millbrae, CA: The California Academic Press.
- Facione, P. (2011). *Think critically*. Upper Saddle River, NJ: Prentice Hall.
- Facione, P., & Facione, N. (1994). *Holistic critical thinking scoring rubric*. Millbrae, CA: The California Academic Press.
- Facione, P., & Gittens, C. A. (2015). *Think critically*. New York, NY: Pearson Education.
- Fahim, M., & Hashtroodi, P. (2012). The effect of critical thinking on developing argumentative essays by Iranian EFL university students. *Journal of Language Teaching & Research*, 3(4), 632-638. <http://dx.doi.org/10.4304/jltr.3.4.632-638>
- Furedy, C., & Furedy, J. (1985). Critical thinking: Toward research and dialogue. *New Directions for Teaching and Learning*, 1985(23), 51-69. <https://doi.org/10.1002/tl.37219852307>
- Goatly, A. (2000). *Critical reading and writing: an introductory coursebook*. London: Routledge.
- Guiller, J., Durndell, A., & Ross, A. (2008). Peer interaction and critical thinking: face-to-face or online discussion? *Learning and Instruction*, 18(2), 187-200. <https://doi.org/10.1016/j.learninstruc.2007.03.001>
- Hanson, T. L. (2004). An oral storytelling approach to teaching writing in the primary grades. *Storytelling, Self, Society*, 1(1), 74-91. <https://doi.org/10.1080/15505340409490259>
- Hedgcock, J., & Lefkowitz, N. (1994). Feedback on feedback: Assessing learners' receptivity to teacher response in L2 composing. *Journal of Second Language Writing*, 3, 141-163. [https://doi.org/10.1016/1060-3743\(94\)90012-4](https://doi.org/10.1016/1060-3743(94)90012-4)
- Hounsell, D. (1984). Essay planning and essay writing. *Higher Education Research and Development*, 3(1), 13-31. <https://doi.org/10.1080/0729436840030102>
- Hu, M. (2017). Teachers' intervention in developing English majors' critical thinking when teaching writing in Chinese universities. *Theory and Practice in Language Studies*, 7(12), 1290-1294. <http://dx.doi.org/10.17507/tpls.0712.16>
- Jones, F. R. (1995). Learning an alien lexicon: A teach-yourself case study. *Second Language Research*, 11(2), 95-111. <https://doi.org/10.1177/026765839501100202>
- Ku, K. Y. (2009). Assessing students' critical thinking performance: Urging for measurements using multi-response format. *Thinking skills and creativity*, 4(1), 70-76. <https://doi.org/10.1016/j.tsc.2009.02.001>
- Landis, M., Swain, K. D., Friehe, M. J., & Coufal, K. L. (2007). Evaluating critical thinking in class and online: Comparison of the Newman method and the Facione rubric. *Communication Disorders Quarterly*, 28(3), 135-143. <https://doi.org/10.1177/15257401070280030301>
- Liu, F., & Stapleton, P. (2018). Connecting writing assessment with critical thinking: An exploratory study of alternative rhetorical functions and objects of enquiry in writing prompts. *Assessing Writing*, 38, 10-20. <https://doi.org/10.1016/j.asw.2018.09.001>
- McKeachie, W. (1994). *Teaching tips: strategies, research, and theory for college and university teachers*. Lexington, MA: DC Heath & Continuing Education.

- Mojica-Díaz, C. C., & Sánchez-López, L. (2010). Constructivist grammatical learning: a proposal for advanced grammatical analysis for college foreign language students. *Foreign Language Annals*, 43(3), 470-487. <https://doi.org/10.1111/j.1944-9720.2010.01093.x>
- Niu, L., Behar-Horenstein, L. S., & Garvan, C. W. (2013). Do instructional interventions influence college students' critical thinking skills? A meta-analysis. *Educational Research Review*, 9, 114-128. <https://doi.org/10.1016/j.edurev.2012.12.002>
- Oshima, A., & Hogue, A. (2007). *Introduction to academic writing*. White Plains, NY: Pearson Longman.
- Plath, D., English, B., Connors, L., & Beveridge, A. (1999). Evaluating the outcomes of intensive critical thinking instruction for social work students. *Social Work Education*, 18(2), 207-217. <https://doi.org/10.1080/02615479911220201>
- Quitadamo, I., & Kurtz, M. (2007). Learning to improve: using writing to increase critical thinking performance in general education biology. *CBE-Life Sciences Education*, 6(2), 140-154. <https://doi.org/10.1187/cbe.06-11-0203>
- Reed, J. H., & Kromrey, J. D. (2001). Teaching critical thinking in a community college history course: Empirical evidence from infusing Paul's model. *College Student Journal*, 35(2), 201-201.
- Schwartz, P. W. (1968). A compendium of methods for the teaching of science to gifted children. *Science Education*, 52(2), 130-138. <https://doi.org/10.1002/sce.3730520206>
- Watson, G., Glaser, E. (1994). *Watson–Glaser critical thinking appraisal manual*. New York, NY: Harcourt, Brace & World.
- Willingham, D. T. (2008). Critical thinking: Why is it so hard to teach? *Arts Education Policy Review*, 109(4), 21-32. <https://doi.org/10.3200/AEPR.109.4.21-32>
- WN, U. D. A., Syahri, I., & Simaibang, B. (2018). The correlations among critical thinking, linguistic intelligence, and writing ability of the sixth semester students of FKIP PGRI University of Palembang. *ELTE Journal*, 6(1), 54-67. <http://dx.doi.org/10.31851/elte.v6i1.1555>