Pertanika J. Soc. Sci. & Hum. 26 (4): 2229 - 2242 (2018)

PERTANIKA JOURNALS

SOCIAL SCIENCES & HUMANITIES

Journal homepage: http://www.pertanika.upm.edu.my/

Assessing Omani University Entrants' Critical Thinking Skills with the Cornell Class-Reasoning Test Form X

Sandhya Rao Mehta[1*], Rahma Al-Mahrooqi[2], Christopher Denman[3] and Khalsa Al-Aghbari[4]

[1]Department of English Language and Literature, College of Arts and Social Sciences, Sultan Qaboos University, PO Box 42, Postal Code 123, Al Khodh, Oman
[2]Deputy Vice-Chancellor of Postgraduate Studies and Research, Sultan Qaboos University, PO Box 17, Postal Code 123, Al Khodh, Oman
[3]Office of the Deputy Vice-Chancellor of Postgraduate Studies and Research, Sultan Qaboos University, PO Box 17, Postal Code 123, Al Khodh, Oman
[4]Department of English Language and Literature, College of Arts and Social Sciences, Sultan Qaboos University, PO Box 42 PC 123, Oman

ABSTRACT

This study examined the critical thinking skills of first year students at a public university in Oman to determine whether these skills were adequately developed at the school level. An adapted version of the Cornell Class-Reasoning Test Form X was administered to 60 students who had just graduated high school and entered the university. Descriptive analysis of test results revealed that participants had failed to master four of the five principles assessed by the test's item groups, while receiving scores that suggested neither mastery nor failure to master the remaining principle. The overall average correct percentage for the sample of 45.8% was comparable to results from grade 4 students in the United States offered by the test creators. Independent samples *t*-tests indicated statistically significant differences on overall test scores based on gender and level of study in the English foundation program, although no differences based on age were found. The paper concludes by arguing that critical thinking needs to be better integrated into the curriculum in Oman's education system in order to adequately prepare learners for the demands of university study and the workforce.

ARTICLE INFO

Article history:
Received: 25 July 2017
Accepted: 12 April 2018
Published: 24 December 2018

E-mail addresses:
rao@squ.edu.om (Sandhya Rao Mehta)
mrahma@squ.edu.om (Rahma Al-Mahrooqi)
denman@squ.edu.om (Christopher Denman)
aghbari222@hotmail.com (Khalsa Al-Aghbari)
* Corresponding author

ISSN: 0128-7702
e-ISSN 2231-8534

Keywords: Critical thinking, Cornell Class-Reasoning Test, EFL, Oman

INTRODUCTION

Critical thinking has recently become a more prominent objective of education systems across levels worldwide. Whether defined as a skill or a disposition, critical thinking is invariably an important component of all disciplines, necessitating its inclusion in various curricula. This is even more important in the current social and economic environment where graduates typically change jobs more often than in the past and where transferable skills remain paramount. Schools thus need to be able to make critical thinking a significant component of the curriculum, whether this is through integrating it with various disciplines, teaching it as a separate skill, or a combination of approaches.

Due to the importance of critical thinking skills for students' future academic and professional careers, Oman's Basic Education system places a heavy focus on developing these skills. Despite this emphasis, research from the sultanate suggests that Omani learners' critical thinking skills are not adequately developed during their high school studies, with a number of causes including learner attitude, teachers, the curriculum, and assessments, all being offered as potential contributors to this situation (AlKhoudary, 2015; Mehta & Al-Mahrooqi, 2015; Thakur & Al-Mahrooqi, 2015).

However, as of yet very little research from Oman has explicitly focused on the extent to which high school students in the country have developed their critical thinking skills. This study sought to address this issue by administering the Cornell Class-Reasoning Test Form X to 60 Omani students who had recently graduated high school and entered the English preparatory program of Oman's only public university. Results were compared to benchmarks for different school grade levels offered by Ennis, Gardiner, Morrow, Paulus and Ringel (1964), and by McLellan's (2009) university students in the UAE who took the Cornell Conditional-Reasoning Test Form X, which was also developed by Ennis et al. Differences in test results based on gender, foundation program level, and age were also explored.

LITERATURE REVIEW

A central problem surrounding critical thinking is that there is no consensus about its meaning, scope, or application. Introduced into education in the early twentieth century by scholars such as John Dewey, critical thinking became increasingly central to the study of education from the middle of the century through the works of Norris and Ennis (1989), Facione (1990), and Halpern (1998). Ennis (1987) defined critical thinking as "reasonable reflective thinking focused on deciding what to believe or do", while Norris and Ennis (1989) suggested that a critical thinking test should cover the various areas that employed inferences and the result of these inferences that could be included under the rubrics of induction and deduction.

Whether seen as identifying arguments and counter arguments or the skill of deducting from given information, the importance of critical thinking remains central to education at all levels. In response to this call, Fisher (1998) stated that creating a curriculum which placed "the development of thinking skills at the heart of the educational process" should be a goal of education systems, while Elder (2005) noted that "critical thinking is foundational to the effective teaching of any subject". In addition, Atkinson and Ramanathan (1995, cited in Willingham, 2008) focused on the cultural aspects of critical thinking and suggested that "critical thinking is not a skill. There is not a set of critical thinking skills that can be acquired and deployed regardless of context".

However, it may be defined and applied, critical thinking remains pivotal to educational outcomes and objectives in institutions across the globe. This is clear from the vision statements of a large number of high school boards that often claim to develop "inquiring" students, to tertiary educational institutions whose missions include developing high-level problem solving and critical thinking skills. Clearly, it is imperative that students possess such skills as reasoning, deduction, and evaluation before being absorbed into higher education institutions where these skills will increasingly be called upon. While American universities determine and evaluate such skills through tests including the SAT and ACT, similar standardized tests are not as commonly applied in many other countries around the world. As such, various other tests, such as the Ennis-Weir Critical Thinking Essay Test, the Watson-Glaser Critical Thinking Test, and the California Critical Thinking Test, are used to determine the levels of reasoning and application of information by students in various disciplines.

The Use of Standardized Tests

The use of standardized tests has itself been subject to much discussion in educational contexts. Rhodes (2011) and Possin (2008) suggested that portfolios were more indicative of student performance, while Benjamin (2014) stated that tests outside a specific content area or discipline, and therefore without context, could not be representative of student capacity for thinking. Liu, Frankel, and Crotts (2014) maintained that motivation was an important consideration in measuring critical thinking, as students who were taking a test whose results did not immediately impact them tended to not do as well as those who were taking a test for college admissions, as a contribution to their course mark, and so on. Hatcher (2011) suggested that scores were also dependent on the test that was taken by students and the way learners were trained to take it. Regardless of these concerns, throughout much of the world standardized tests in one form or the other are commonly given to students upon exit from school or on entry to tertiary institutions in order to determine student achievement outside their GPA scores (Benjamin, 2014; Hatcher, 2006; Niu, Behar-Horenstein, & Garvan, 2013).

The use of standardized tests in the Middle East and North Africa (MENA) region has been the subject of increasing research activity as it is only in the past 30 years that educational institutions have been established in many of the region's nations. Results from studies in Jordan (Bataineh & Zghoul, 2006), Iran (Azin & Tabrizi, 2016), Turkey (Sahin, 2016), the UAE (McLellan, 2009), and Oman (AlKhoudary, 2015; Mehta & Al-Mahrooqi, 2015; Thakur & Al-Mahrooqi, 2015) have suggested that students in high schools and universities in the region need much more training in critical thinking as part of both the school curriculum and in their specialized college disciplines. However, these studies have used a variety of tests across different disciplines and educational levels, making it difficult to arrive at a more definitive assessment of the role of critical thinking in educational institutions in the MENA region. This study, therefore, enlarges the scope of existing research by including first year students enrolled in a public university in Oman in order to measure their critical thinking skills.

The Cornell Class-Reasoning Test Form X (Ennis et al., 1964) is part of a group of tests known as the Cornell Critical Thinking Tests, which also include the Cornell Conditional-Reasoning Test (Methodology). A number of researchers have used one or more of these tests to assess students' critical thinking skills in a variety of nations. For example, in the United States, Hughes (1992) reported only weak correlations of 0.15–0.17 between the Cornell Critical Thinking Tests and student grades, although correlations with scholastic aptitude and intelligence measures were 0.50. Nolan and Brandon (1984) used a similar study to investigate the deductive reasoning skills of high school students in Jamaica and found that, while gender and academic achievement in the form GPA were not indicators of higher test scores, streams into which students entered did have an impact of their scores with science students scoring higher than most others.

In the MENA region, Bataineh and Zghoul's (2006) study using the Cornell Critical Thinking Test Level Z with master's students in Jordan indicated that learners "performed quite poorly", although the authors reported that older males and younger females performed better on the test. In the UAE, McLellan (2009) investigated 361 business students using the Cornell Conditional-Reasoning Test Form X, which, like the current instrument, was designed by Ennis et al., (1964) for school-level learners. Based on Nolan and Brandon's (1984) work, the author hypothesized that there would be no difference in performance based on gender, that older students would perform better than younger ones, that GPA would be a significant indicator of test scores, and that time spent in the business program would determine test performance. The author reported an overall mean score of 45.93 (SD = 11.92) from 72 questions (63.8%) as compared to an overall mean score of 56.6 (SD = 14.00) among 17-year-old students in the United States (78.6%).

McLellan (2009) also reported that while gender was not a significant indicator of critical thinking skills, older students performed better than younger ones. However, no relationship was found between GPA or the number of credits completed by students, contrary to the hypotheses put forward. McLellan concludes by suggesting that the Cornell Conditional-Reasoning Test Form X could be used for further analysis of critical thinking skills in the UAE. The current research attempts a similar approach in Oman though with the use of the Cornell Class-Reasoning Test Form X.

METHODOLOGY

Research Site and Sample

Sultan Qaboos University is the only fully publicly funded university in Oman. It was established as the first university in Oman in 1986 to enable citizens from around the sultanate to have access to tertiary education. Students are chosen for entry on the basis of grades received in the *Thanawiya Amma,* the high school diploma, as well as other diplomas offered by private schools (such as the IGCSE or the IB). After acceptance into the university, all non-exempt students are required to do foundation courses in English language, IT and mathematics, in order to prepare them for their specializations. Students who present an overall IELTS score of 5.0 or above are exempt from the English component and can enter their colleges directly.

Research Question and Objectives

The present study investigated the following questions:

What is the level of Critical Thinking among college entrants in Oman?

What impact, if any, do the variables of gender, English foundation level, and age have on participants' Cornell Critical Thinking test results?

In order to test the critical thinking skills of these students and, therefore, gain a clearer idea of whether critical thinking skills are being adequately developed at the school level, an adapted version of the Cornell Critical Thinking Test Form X was given to 60 foundation students at the research site. To recruit participants, students in one Level 4 (intermediate) and one Level 6 (upper intermediate) English class were approached to sit the test during class time. After being reminded of the study's voluntary and anonymous nature, and that their choice whether to participate or not would not have any impact on their standing in the class or the university, a total of 60 students agreed to complete the test. 40.54% of participants were male and 59.46% female, with 54.05% studying at Level 4 and the remaining 45.95% in Level 6. All participants were either 18 or younger (83.78%) or 19 or older (16.22%), and came from governorates across the country, including Al-Dakhilia (21.62%), Al-Batinah South (18.92%), Muscat (13.51%), Al-Dhahira (13.51%), Al-Sharqiya North (13.51%), Al-Batinah North (8.11%), Al-Sharqiya South (5.41%), Musandam (2.70%), and Al-Buraimi

(2.70%). Participants were administered the test during their foundation English classes by their regular instructors, and did not receive any explicit test training beforehand.

Instrument

The biggest challenge in choosing a critical thinking test for the current research context included the EFL status of participants and issues associated with conflating English ability with critical thinking skills for tests administered in English on the one hand, and issues associated with seeking to accurately translate a test from English to Arabic on the other. In order to address these concerns, an adapted English-language version of the Cornell Class-Reasoning Test Form X (Ennis et al., 1964) was employed as the test presents a series of critical thinking questions in simplified language that is even appropriate, according to the test creators, for primary-school children. Each question is introduced by a supposition followed by the question "Then would this be true?" accompanying a statement. For example, question 6 of the original Cornell Class-Reasoning Test Form X is worded and formatted in the following way:

6. Suppose you know that
   X is next to Y
Then would this be true?
   Y is next to X.

Test takers must choose one of three answers to each question: Yes, No, or Maybe. Yes indicates that the final statement must be true, No that the statement cannot be true, and Maybe that it may or may not be true but that there was not enough information provided to be sure.

While the language, formatting, and layout of the test's questions were considered to be largely appropriate for participants, some questions did, nonetheless, contain cultural references and/or Western names that were changed to make the test more culturally-appropriate. Examples include the change of names in one of the example questions from "Bill is next to Sam" to "Mazin is next to Ahmed", and, "All the cars in the garage are Mr. Smith's" to "All the cars near the house are Saif's" (private garages are not as common in Oman as they are in many Western nations). These revisions were made during discussions with a panel of five instructors in the fields of education and linguistics at the research site which included members of the research team.

Revisions also took into account time constraints on data collection. As a result, the number of questions included in the test was reduced from the original 72 across 12 item groups to 36 questions across six items groups of six questions each (a course of action recommended by Ennis et al., 1964) in addition to the three example questions. These 36 questions cover all three content components of the test with the same frequency as the full-length test. These content components are concrete familiar, symbolic, and suggestive. Concrete familiar refers to questions about concrete articles and qualities that test-takers are most likely familiar with; however, these questions refer to a specific example featuring these items that test-takers are not familiar with—for example, "John's car." Symbolic refers to those cases where symbols, such as "X,"

"Y," "A", and so on, are used instead of references to particular objects, while suggestive relates to those questions where content is familiar to test-takers, but the truth or falsity of the content is not known. The items featured in the adapted test also cover five of the principles that were recommended in the manual for younger test takers, with principle 3 featured twice. These are:

Principle 1: All As are Bs.: At least some As are not Bs. All As are Bs. All Bs are Cs.: All As are Cs.

Principle 2: All As are Bs. All Bs are Cs. : All As are Cs.

Principle 3: All As are Bs.: All Bs are As.

Principle 3: All As are Bs. All Cs are Bs. : At least some Cs are As.

Principle 4: No As are Bs.: No Bs are As.

Principle 5: All As are Bs. No Cs are Bs. : At least some As are Cs.

Item groups of six questions each from the adapted version of the test employed here are as follows:

Item group 1 (principle 1): items 5, 32, 26, 13, 19, 36

Item group 2 (principle 2): items 4, 37, 24, 11, 16, 28

Item group 3 (principle 3): items 8, 21, 29, 34, 27, 38

Item group 4 (principle 3): items 6, 10, 23, 15, 31, 20

Item group 5 (principle 4): items 7, 14, 17, 30, 35, 25

Item group 6 (principle 5): items 9, 18, 39, 22, 12, 33

Ennis et al.'s (1964) test–retest reliability calculations of the test's content components with 329 school-level participants in grades 4, 6, 8, 10, and 12 reports an overall mean r value, calculated with the use of Fisher's z, of .83. In terms of individual content components, these values are: concrete familiar (0.79), symbolic (0.50), and suggestive (0.63). Test–retest r values for the six item groups employed in the current study are: Item group 1 (0.60), item group 2 (0.55), item group 3 (0.69), item group 4 (0.52), item group 5 (0.38), and item group 6 (0.52). (The adapted version of the test is available from the researchers upon request.)

Ennis et al. (1964) offer detailed results of school students on the class-reasoning test. The authors applied a formula of (R-W/2)+27 (where R = number of correct responses and W = number of incorrect responses) to calculate a total score from 99 for the 72 test items. Results for Ennis et al.'s student samples are: grade 4 – 48.5%; grade 6 – 57.5%; grade 8 – 58.8%; grade 10 – 74.2%; grade 12 – 73.1%. However, Ennis et al. (1964) state that the main concern of the test is not the overall scores themselves, but rather determining whether test-takers have mastered the principle that item groups are based on. In order to determine this, Ennis and Paulus (1965, cited in Johnson & Posner, 1971) state that mastery is demonstrated when a test-taker records correct answers for 5 or 6 items from an item group, while failure to master a principle occurs when 3 or fewer items have been answered correctly. Four correct answers for

an item group demonstrates neither mastery nor a failure to master the principle. For ease of interpretation in the current study, these numbers have been converted to percentages, with scores of around 83.3% or higher indicating mastery of the principle, scores of around 50.0% or lower indicating failure to master the principle, and score of around 66.7% indicating neither mastery nor failure to master the principle.

In the current study, the marking formula offered by Ennis et al. (1964) has not been applied as it is not applicable to the shortened version of the test. Instead, all participant answers to each item on the critical thinking test were marked as either correct or incorrect, with any missing responses also being marked incorrect. Items were then placed into the aforementioned six item groups. Correct and incorrect percentages for each item and each item group were then calculated. In addition, total correct and incorrect percentages were also calculated for the three test components of concrete familiar, symbolic, and suggestive. Three independent samples *t*-tests were then conducted to examine if statistically significant differences in overall test scores existed based on the variables of gender, English foundation level (Level 4 and Level 6), and age (18 or younger and 19 or older). A Bonferroni correction was made to decrease the possibility of type I error for these *t*-tests, with the resultant *p* value being set at $p \leq 0.03$.

RESULTS

Item group 1 was concerned with principle 1: 'All As are Bs.: At least some As are not Bs. All As are Bs. All Bs are Cs.: All As are Cs' (see Table 1). The overall correct percentage for this item group of 61.7% suggests that participants had neither mastered nor failed to master the principle. This item group recorded amongst the highest correct percentages for individual questions, with almost all of these being above 60%. The one exception to this was for question 36, with only 48.6% of participants recording correct answers.

Table 1
Item group 1

Question Number	Percent Correct	Percent Incorrect
5 (CF)	64.9	35.1
32 (CF)	70.3	29.7
26 (CF)	62.2	37.8
13 (CF)	62.2	37.8
19 (SY)	62.2	37.8
36 (SU)	48.6	51.4
Total	61.7	38.3

Note: CF = concrete familiar; SY = symbolic; SU = suggestive

Table 2 features item group 2, which was associated with principle 2: 'All As are Bs. All Bs are Cs.: All As are Cs'. This item group received an overall correct percentage of 52.3%, thereby suggesting participant failure to master the principle. Although around 70.3% of participants offered the correct answer for question 24, the percentage of correct responses for the remaining 6 items were all under 57%, with this being as low as 35.1% for question 28.

Table 2
Item group 2

Question Number	Percent Correct	Percent Incorrect
4 (CF)	56.8	43.2
37 (CF)	40.5	59.5
24 (CF)	70.3	29.7
11 (CF)	54.1	45.9
16 (SY)	56.8	43.2
28 (SU)	35.1	64.9
Total	52.3	47.7

Table 3
Item group 3

Question Number	Percent Correct	Percent Incorrect
8 (CF)	21.6	78.4
21 (CF)	18.9	81.1
29 (CF)	18.9	81.1
34 (CF)	37.8	62.2
27 (SY)	27.0	73.0
38 (SU)	18.9	81.1
Total	23.8	76.2

Table 3 indicates that item group 3 received the lowest overall percentage of correct responses of 23.8%. This item group was associated with principle 3: 'All As are Bs. All Cs are Bs.: At least some Cs are As'. Participant responses indicated a failure to master the principle, with the highest percentage of correct responses being for question 34 (37.8%) and with the lowest correct percentage of 18.9% being shared by questions 21, 29, and 38.

Item group 4 was associated with the same principle as item group 3 as stated earlier (see Table 4). Participants here again demonstrated a failure to master the principle, with an overall correct percentage of 37.8%. No questions here recorded overall correct percentages of above 50%, with the highest being for question 6 (48.6%).

Table 5 features correct percentages for participants for item group 5 that was associated with principle 4: 'No As are Bs.: No Bs are As'. Correct percentages here again suggest that students failed to master the principle, with the highest correct percent being for question 14 (59.5%). The overall correct percent for this item group was 51.4%.

Table 4
Item group 4

Question Number	Percent Correct	Percent Incorrect
6 (CF)	48.6	51.4
10 (CF)	43.2	56.8
23 (CF)	37.8	62.2
15 (CF)	45.9	54.1
31 (SY)	21.6	78.4
20 (SU)	29.7	70.3
Total	37.8	62.2

Table 5
Item group 5

Question Number	Percent Correct	Percent Incorrect
7 (CF)	48.6	51.4
14 (CF)	59.5	40.5
17 (CF)	54.1	45.9
30 (CF)	45.9	54.1
35 (SY)	51.4	48.6
25 (SU)	48.6	51.4
Total	51.4	48.6

Table 6
Item group

Question Number	Percent Correct	Percent Incorrect
9 (CF)	59.5	40.5
18 (CF)	48.6	51.4
39 (CF)	24.3	75.7
22 (CF)	56.8	43.2
12 (SY)	54.1	45.9
33 (SU)	35.1	64.9
Total	46.4	53.6

Table 6 offers percentages correct for questions from item group 6, which was associated with principle 5: 'All As are Bs. No Cs are Bs.: At least some As are Cs'. The overall correct percentage for this item group was 46.4%. Percentages of correct responses suggest that students failed to master the principle. The lowest level of correct percentage was for question 39 (24.3%) and the highest was for question 9 (59.5%).

The overall correct percentage for the entire test was 45.8%, with correct percentages for the three content components being: symbolic (45.5% across 6 questions), suggestive (36.0% across 6 questions), and concrete familiar (43.2% across 24 questions). Independent samples *t*-tests were conducted to examine if statistically significant differences in test scores existed between gender, English foundation level (Level 4 and 6), and age (18 or younger and 19 or older). As stated above, the accepted *p* value was set at 0.03.

The first *t*-test exploring the variable of gender revealed statistically significant differences in scores across the critical thinking test ($p = 0.002$) with female participants receiving higher scores than their male classmates. Females recorded an average score across the test of 18.36 (51.0%) and males recorded an average score of 13.53 (37.6%). Level 6 students also had significantly higher scores than students in Level 4 ($p = .011$). Here, Level 6 students received overall test scores of 18.53 (51.5%) while Level 4 students recorded an average score across the test of 14.60 (40.6%). No statistically significant differences based on age were found.

DISCUSSION

The integration of critical thinking instruction into the school curriculum is a priority of education systems around the world, including in Oman. Fisher (1998) argued that school curricula should be designed to place critical thinking skills at the heart of the educational process, with critical thinking skills such as induction, reasoning, inferencing, and deduction being integral to textbooks, teaching, assessment, and so on. Research from Oman suggests that Omani students generally do not have well-developed critical thinking skills (AlKhoudary, 2015; Mehta & Al-Mahrooqi, 2015; Thakur & Al-Mahrooqi, 2015), with similar findings being reported from other countries within the MENA region (Azin & Tabrizi, 2016; Bataineh & Zghoul, 2006; McLellan, 2009; Sahin, 2016).

Results from administering the Cornell Class-Reasoning Test Form X to 60 Omani students in this study reveal an overall

correct percentage of 45.8%. Ennis et al., (1964), as the developers of the test, reported overall correct percentages for an American school sample of 48.5% for grade 4, 57.5% for grade 6, 58.8% for grade 8, 74.2% for grade 10, and 73.1% for grade 12. Based on these figures, it appears as though the Omani university participants in the current study received similar results on the test to grade 4 students in the United States around 50 years previously. Of course, a number of concerns must be taken into account when making any claims about this figure, including the non-native English speaker status of the current participants and the fact that the test may not have been viewed as having any practical significance to respondents, which may be reflected in their low scores (Hatcher, 2011). Finally, it is also likely that participants' lack of explicit training on the test negatively affected their scores (Liu, Frankel, & Crotts, 2014).

While few researchers have applied the Cornell Class-Reasoning Test Form X to students in countries in the MENA region, McLellan's (2009) study of university students in the UAE used the Cornell Conditional-Reasoning Test Form X, which was also developed by Ennis et al. (1964). Although these tests take a different approach to assessing test-takers' critical thinking skills,

McLellan's research is referenced here as it one of the few that has adopted a Cornell Critical Thinking Test within a Arab Gulf nation, thereby offering some insight about critical thinking skill development within a similar context. As stated earlier, the author reported an overall correct percentage of 63.8% on the Cornell Conditional-Reasoning Test Form X—around 20% higher than in the current study, again suggesting a limited development of critical thinking skills in the Omani school system.

Results across each of the test's six item groups reveal that participants failed to master four of the five principles tested. These were:

Principle 2: All As are Bs. All Bs are Cs.: All As are Cs. (52.3% correct)

Principle 3: All As are Bs.: All Bs are As. (23.8% correct)

Principle 3: All As are Bs. All Cs are Bs. : At least some Cs are As. (37.8% correct)

Principle 4: No As are Bs.: No Bs are As. (51.4% correct)

Principle 5: All As are Bs. No Cs are Bs. : At least some As are Cs. (46.4% correct)

The principle on which participants received the highest overall correct percent of 61.7% was for Principle 1: 'All As are Bs.: At least some As are not Bs. All As are Bs. All Bs are Cs.: All As are Cs'; however, this level does not suggest mastery of the principle, but rather neither mastery nor failure to master it based on the guidelines offered by Ennis and Paulus (1965). It is also below the approximately 71.7% correct percentage recorded by Ennis et al.'s (1964) grade 4 participants. A similar pattern is observable when comparing overall correct percentages for the test's three content components of symbolic, suggestive, and concrete familiar. Omani respondents here

received overall correct percentages that were comparable with Ennis et al.'s grade 4 participants for concrete familiar (43.2% for the current study versus 53.1% in Ennis et al., 1964) and suggestive (36.0% versus 37.5%) questions, although were closer in results to Ennis et al.'s, grade 6 and grade 8 students for symbolic questions (45.5% versus 45.8% for grade 6 and 44.2% for grade 8).

Independent samples *t*-tests were conducted to explore the relationship between the variables of gender, level in the English foundation program, and age on the overall test score. Female participants scored around 13% higher on the Cornell Class-Reasoning Test Form X than their male colleagues, while Level 6 students scored around 11% higher than students in Level 4 of the English foundation program. Differences based on age were not found. The first of these findings adds to the rather complex picture in the literature, in which gender was not found to be related to test results on Cornell Class-Reasoning and Conditional-Reasoning Tests by Nolan and Brandon (1984) and McLellan (2009), but was found to be significant by Bataineh and Zghoul (2006) in Jordan. In addition, both Bataineh and Zghoul, and McLellan reported statistically significant differences on test results based on age, while the current study did not find any such differences.

There are a number of limitations, in addition to those discussed earlier, which must be taken into account when interpreting these results. First, the small sample size of 60 students means that the external validity of these findings to the population under study—Omani school students—is necessarily limited. Similarly, assessing high school graduates who have gained entry into the most prestigious university in the country means that these respondents may be substantially different in critical thinking skills than those who did not enter the university, who went to different universities and colleges, or who pursued professional careers after graduating. As a result, future research should seek to assess students' critical thinking skills while they are still in high school, rather than waiting until they have been selected for higher level study. In addition, the modified nature of the test may also have reduced its reliability and validity, even if these modifications were done in line with Ennis et al.'s (1964) guidelines. Finally, the status of the participants as non-native speakers of English, and the relative lack of studies that offer regional benchmarks for the Arab Gulf nations, may also have impacted the results and their interpretation here.

CONCLUSION

This study examined the critical thinking skills of first year university students in Oman in order to determine whether these skills are being effectively taught at the school level. In order to achieve this, the Cornell Class-Reasoning Test Form X was administered to 60 students enrolled in the English language component of a foundation program in Oman's only public university. Results revealed that participants had failed to master four of the five principles tested,

and that they also scored at levels that were mostly comparable to grades 4 to 6 school learners in the United States as reported by the test developers (Ennis et al., 1964). Although the Cornell Class-Reasoning Test Form X has not been widely used with similar populations in the Arab Gulf region previously, participants in this study received lower scores than a similar group of students at an Arab Gulf university who sat the Cornell Class-Conditioning Test Form X (McLellan, 2009), which was also developed by Ennis et al. (1964). While statistically significant differences were found based on the variables of gender and English foundation program level, Omani students nonetheless appear to generally struggle with critical thinking skills, at least as measured by the current test. These findings suggest that the way critical thinking is integrated into the curriculum in Omani schools may need to be reviewed, with particular attention paid to such areas as the curriculum itself, teaching practices, textbooks, assessment, and so on. Such a course of action may be found by future research to be necessary in order to help prepare school learners in Oman for their future academic and professional careers.

REFERENCES

AlKhoudary, Y. A. M. (2015). The effect of teaching critical thinking on Al-Buraimi university college students' writing skills: A case study. *International Journal of Applied Linguistics & English Literature, 4*(6), 212-219. Retrieved July 25, 2017, from http://www.journals.aiac.org.au/index.php/IJALEL/article/view/1798/1672

Azin, N., & Tabrizi, H. H. (2016). The relationship between critical thinking ability of Iranian English translation students and their translation ability. *Theory and Practice in Language Studies*, *6*(3), 541–548. doi: 10.17507/tpls.0603.12

Bataineh, R. F., & Zghoul, L. (2006). Jordanian TEFL graduate Students' use of critical thinking skills (as measured by the Cornell Critical Thinking Test, Level Z). *International Journal of Bilingual Education and Bilingualism, 9*(1), 33–50. doi: 10.1080/13670050608668629

Benjamin, R. (2014). Two questions about critical-thinking tests in higher education. *Change: The Magazine of Higher Learning*, *46*(2), 24–31. doi: 10.1080/00091383.2014.897179

Hatcher, D. L. (2006). Stand-alone versus integrated critical thinking courses. *The Journal of General Education*, *55*(3/4), 247–272.

Elder, L. (2005). Critical Thinking as the key to the learning college: A professional development model. *New Directions for Community Colleges*, *130*, 39–48.

Ennis, R. H. (1987). A taxonomy of critical thinking dispositions and abilities. In J. B. Baron & R. J. Sternberg (Eds.), *Teaching thinking skills: Theory and practice* (pp. 9–26). New York, USA: W.H. Freeman.

Ennis, R. H. & D. H. Paulus. (1965). *Critical thinking readiness in grades 1-12 (Phase I: Deductive logic in adolescence)*. Ithaca, NY: Cornell University.

Ennis, R. H., Gardiner, W. L., Morrow, R., Paulus, D., & Ringel, L. (1964). *The Cornell class reasoning test, form X. Cornell critical thinking test series.* Retrieved July 25, 2017, from http://faculty.education.illinois.edu/rhennis/cornellclassreas.pdf

Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction*. The Delphi Report

(ED315-423). Retrieved July 25, 2017, from https://philarchive.org/archive/FACCTA

Fisher, R. (1998). Thinking about thinking: Developing metacognition in children. *Early Child Development and Care*, *141*, 1–13.

Halpern, D. F. (1998). Teaching critical thinking for transfer across domains: Dispositions, skills, structure training, and metacognitive monitoring. *American Psychologist*, *53*(4), 449–455.

Hatcher, D. L. (2011). Which test? Whose scores? Comparing standardized critical thinking tests. *New Directions for Institutional Research*, *149,* 29–39. doi: 10.1002/ir.378

Hughes, J. N. (1992). Review of the Cornell critical thinking tests. In J. J. Kramer & J. C. Conoley (Eds.), *The eleventh mental measurements yearbook*. Lincoln, USA: Buros Institute of Mental Measurements.

Johnson, M., & Posner, G. J. (1971). *Testing the effect of verbal-quantitative aptitude discrepancy on the learning of deductive reasoning through programed instruction*. Albany, USA: State University of New York at Albany.

Liu, O. L., Frankel, L., & Crotts, R. K. (2014). Assessing critical thinking in higher education: Current state and directions for next-generation assessment. *ETS Research Reports Series, 1*, 1–23. doi: 10.1002/ets2.12009

McLellan, J. D. (2009). Assessing the critical thinking abilities of United Arab Emirates business students. *Learning and Teaching in Higher Education: Gulf Perspectives, 6*(2). Retrieved July 25, 2017, from https://ssrn.com/abstract=1932038

Mehta, S., & Al-Mahrooqi, R. (2015). Can thinking be taught? Linking critical thinking and writing in an EFL context? *RELC Journal*, *46*(1) 23–36.

Niu, L., Behar-Horenstein, L. S., & Garvan, C. (2013). Do instructional interventions influence college students' critical thinking skills? A meta-analysis. *Educational Research Review,* 9, 114–128. doi: 101016/jedurev201212002

Nolan, S. P., & Brandon, E. P. (1984). *Conditional reasoning in Jamaica.* Retrieved July 25, 2017, from https://files.eric.ed.gov/fulltext/ED260974.pdf

Norris, S. P., & Ennis, R. H. (1989). *Evaluating critical thinking*. Pacific Grove, USA: Midwest Publications.

Possin, K. (2008). A field guide to critical-thinking assessment. *Teaching Philosophy, 31*(3), 201–228.

Rhodes, T. L. (2011). Making learning visible and meaningful through electronic portfolios. *Change: The Magazine of Higher Learning, 43(1), 6-13. doi: 10.1080/00091383.2011.538636*

Sahin, U. (2016). The evaluation of education faculty students' critical thinking skills. *International Journal of Social Sciences and Education*, *6*(2), 133–140.

Thakur, V., & Al-Mahrooqi, R. (2015). Orienting ESL/EFL students towards critical thinking through pictorial inferences and elucidation: A fruitful pedagogic approach. *English Language Teaching*, *8*(2), 128–133.

Willingham, D. T. (2008). Critical thinking: Why is it so hard to teach? *Arts Education Policy Review, 109(4), 21-32.*