

QUANTIFICATION OF CRITICAL THINKING SKILLS AFTER WITH COMPUTER QUIZ GAMES IN AN INTRODUCTORY SCIENCE COURSE

Szu Szu Ling^{1,2}, Fabrice Saffre³, Deborah Gater^{1,4}, Lilia Halim⁵, & Abdel F. Isakovic^{6,7}

¹Department of Chemistry, Khalifa University of Science and Technology, Abu Dhabi (UAE)

²Department of Preparatory Program, Khalifa University of Science and Technology, Abu Dhabi (UAE)

³British Telecom Research and Innovation, Ipswich (United Kingdom)

⁴University College, London (United Kingdom)

⁵Science Department, Faculty of Education, National University of Malaysia, Bangi (Malaysia)

⁶Department of Physics, Khalifa University of Science and Technology, Abu Dhabi, (UAE)

⁷Cornell University, Ithaca, NY (USA)

Abstract

Critical thinking skills, in general, and specifically, within the framework of introductory science courses in engineering and science majors, are increasingly a focus of education reforms. This prioritized goal comes along with the ongoing efforts to phase out traditional teaching and steadily replace it with interactive engagement learning and teaching methods. To respond to such goals without being disruptive in the traditional teaching and learning process, we have developed computer quiz games (CQG), implemented for the purpose of enhancing students' interactive engagement and learning gains in Introductory General Chemistry for Engineering Majors in the English as a Second Language (ESL) environment. The effectiveness of CQGs was quantified using the post-test vs. pre-test approach via FECAT (Freshman Engineering Chemistry Assessment Test) test with Hake's Gain between 15% and 35%. In the second phase of the analysis, we relied on the mapping of distinct categories of FECAT questions to Bloom taxonomy and we classified all FECAT questions into Higher Order Thinking (HOT) and Lower Order Thinking (LOT) skills. We present how HOTs and LOTs get modified and, to a varied degree, improved, with the application of computer quiz games. The results of the application of CQGs were also compared with the application of the paper-based quizzes containing the same questions, offering an insight into the influence of simple computer games on students' hierarchical critical thinking skills.

Keywords: *Critical thinking skills, game-based learning, computer-based learning, higher order thinking.*

1. Introduction

Effective teaching and active learning are essential for college graduates on their path of taking place in today's workforce. Related to active learning is the need to develop, use and sharpen critical thinking skills, along with the ability to solve complex problems (Gokhale, 1995; Scott, 2009). This has led to identifying critical thinking as one of the goals for the academic programs in the higher learning institutions (Van der Werff, 2016, Donald, 2002, McDonald *et al.*, 2014). One of the approaches in quantifying critical thinking skills is to divide the tasks in front of students into lower order thinking (LOT) and higher order thinking (HOT) skills. Ramos *et al.* (2013) discuss how students' level of higher order thinking (HOT) skills affect their performance. Our interest is in elementary, first-year college chemistry course, and note that the work in chemistry has been preceded by Rodrigues and Oliveira (2008), and Tiruneh *et al.*, 2017, who pointed out that the level of critical thinking could be a predictor of the students' performance in physics.

Factors contributing to the problem of learning introductory chemistry are a low level of preparedness from secondary education, low motivation and limitations of the traditional method of instruction students have likely being exposed to before and during college. In Khalifa University (KUST), the Chemistry Department has concerns about the performance and the failure rate in the first-year chemistry course offered to all engineering and science majors. Interviews (Ling *et al.*, 2015) with KUST students and chemistry faculty suggest that students have low motivation and interest to learn chemistry in a traditional classroom and few activities available to them to pull them away from teacher-centered classroom activities. Ideally, one would respond to such needs by developing a fully interactive engagement based course. Given the internal institutional obstacles, such as lack of administrative and technical support to develop and implement fully interactive engagement course, the best alternative was to develop teaching

and learning tools that can augment the ongoing traditional teaching. After carefully examining other efforts attempted at KUST, and in the country and the region (UAE, Gulf Arab countries), such as Collaborative Workshop Physics (Hitt *et al.*, 2014), and Problem Based Learning (Balawi *et al.*, 2015), we opted for the activity that will have the following features:

- (a) Minimally disruptive to ongoing traditional chemistry lectures;
- (b) Supportive of the need to promote learning of the concepts, as opposed to rote-based learning, which often accompanies passive attendance of traditionally delivered lectures;
- (c) Implementable on contemporary wireless digital communication devices, such as: laptops, tablets and smartphones;
- (d) Can be presented as a game-based learning activity.

The preceding design criteria led us to develop computer-based quiz games (CQGs). The details of the CQG activity are broadly compatible with all four criteria listed above. Furthermore, going through the CQG activities on weekly basis enhances critical thinking skills (Kapp, 2012), given that our CQGs satisfy almost all gaming criteria (the presence of the system, one or more players, contains a challenge, follows fixed rules that a player needs to learn to best the game, it is interactive, contains feedback in the form of quantifiable outcome and may have emotional reaction).

2. Objective

The objective of this report is to investigate the effect of CQGs on students' performance with respect to the students' critical thinking skills (HOT and LOT) through a chemistry achievement test called Freshman Engineering Chemistry Assessment Test (FECAT), previously elaborated on (Ling, 2018).

3. Methodology

The present study applied a true experimental design called "Pretest-Posttest" (Gribbons & Herma, 1997; Trochim, 2001) to compare students in treatment (traditional lecture teaching with CQGs) and comparison groups. The comparison (a form of control) is in the form of paper-based quizzes, PQ, where the questions on the paper quiz are the same as those on CQG, but not subject to steps of the gamification process, such as digital interactive environment, or ability choose questions on an individual basis). All the students from each class are randomly assigned to control and treatment groups. Randomization controls all the possible extraneous variables (Adesoji & Babatunde, 2008; Campbell & Stanley, 1963) and enhances validity for possible statistical significance tests.

More specifically, the content of FECAT has been discussed elsewhere (Ling, 2018). We show some example questions later here, in Fig. 2, in order to illustrate the level of difficulty.

3.1. Participants

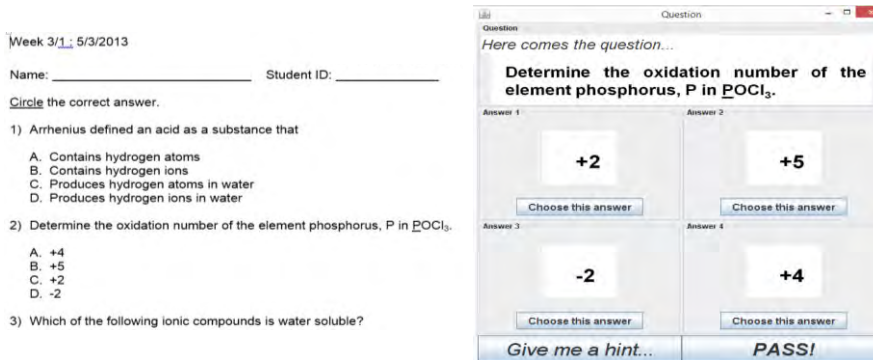
The total of 121 students (63 students for TG and 58 students for CG) participated in this study. In terms of gender, the respondents in TG consisted of 36 (57.1%) male and 27 (42.9%) female students. As for the CG, there were 37 (63.8%) male and 21 (36.2%) female students. We put some effort in making both groups equal in size at the beginning but some students are "disqualified" during or at the end of the semester because either they do not take a sufficient number of CQGs or they do not take the posttest.

3.2. Treatment and control

Each Computer Quiz Game (CQG) has nine questions in total, and students are allowed to choose 5 of these questions, one at the time. The text (the content, often a combination of text and either chemical image or a formula or a chemical reaction) of each question is revealed only after the student selects it. The initial guidance on what to select is only based on the fact that there are three questions of each category (three "easy", three "medium", and three "hard").

The point values per category are 100 points, 200 points, and 300 points. Beyond a fundamental need to have a constraint in a game-like environment, the idea behind limiting the number of choices is to motivate students to rely on their self-assessment of what kind of questions are they comfortable answering, thereby motivating them to strategize during the game and practice various types of questions ahead of the game; the latter activity clearly meant to be a learning opportunity. The paper quiz contains 5 questions only (so, no choosing is possible), but each week we prepared at least 5-6 different versions of the paper quiz (version A - 1 easy, 3 medium, 1 hard; version B - 2 easy, 2 medium, 1 hard; version C - 1 easy, 1 medium, 3 hard, etc.) This way, a form of randomization is present although no student choice is possible.

Figure 1. (Left) a screenshot of the paper quiz; (right) a screenshot of one question of computer game quiz.



3.3. Research instruments

Students' level of thinking skills were divided into higher-order thinking skills (HOT) and lower-order thinking skills (LOT) and assessed through a chemistry achievement test (called Freshman Engineering Chemistry Assessment Test), FECAT by using the classification of Bloom taxonomy. The LOT skills are measured from the questions related to knowledge as well as comprehension and HOT skills (require critical thinking) are measured from the questions related to application, analysis, synthesis and evaluation. There are 15 LOT and 21 HOT questions. We wanted to assess students' HOT skills (more questions) rather than LOT skills (fewer questions). FECAT was also administered at the end of the semester, so it serves as the research instrument in the assessment of learning gain. In order to design a suitable achievement test to address the research questions, issues of validity and reliability have been considered, and addressed at the tie of initial data gathering and analysis (Ling, 2018).

Figure 2. Examples of LOT and HOT questions from FECAT. A full FECAT test available upon request.

EXAMPLES OF LOT QUESTIONS	EXAMPLES OF HOT QUESTIONS
<p>4) Which of these is/are the correct unit for the pressure?</p> <p>A. Pascal (Pa) or Newton/meter²</p> <p>B. Standard atmosphere (atm)</p> <p>C. Millimeters of mercury column (mm Hg)</p> <p>D. Bar</p> <p>E. All of these</p> <p>5) Atoms:</p> <p>A. Exist only in solid materials</p> <p>B. Are made up of protons, electrons and neutrons</p> <p>C. Were considered to be the smallest things in the world</p> <p>D. Can be seen with a microscope</p> <p>E. Were discovered during World War II</p> <p>11) Pure substances:</p> <p>A. Rarely exist around us</p> <p>B. Are homogeneous and have unchanging chemical compositions</p> <p>C. Are heterogeneous and have unchanging chemical compositions</p> <p>D. Only exist in cubic shape</p> <p>E. Mostly exist in spherical shape</p>	<p>7) What natural phenomenon is the most likely demonstration of a chemical reaction that gives out heat?</p> <p>A. Earthquake</p> <p>B. Water freezing</p> <p>C. Fire</p> <p>D. Bacteria and viruses</p> <p>E. Sun</p> <p>8) A compound is found to have a molar mass of 90 g/mol and simplest formula of C₂H₅O. The molecular formula of the substance is:</p> <p>A. C₃H₆O₃</p> <p>B. C₄H₁₀O</p> <p>C. C₆H₁₀O₂</p> <p>D. C₃H₁₄O</p> <p>E. C₄H₁₀O₄</p> <p>9) Which of the ionic compounds below is NOT likely to be soluble in water?</p> <p>A. Calcium carbonate</p> <p>B. Potassium hydroxide</p> <p>C. Ammonium chloride</p> <p>D. Magnesium sulfate</p> <p>E. Lead chloride</p> <p>16) Which of the following is true about sodium hydroxide?</p> <p>A. It is not a corrosive substance</p> <p>B. It is a base that is soluble in water</p> <p>C. It is strong alkali with a low pH value</p> <p>D. It is a weak alkali with a high degree dissociation</p> <p>E. It is strong alkali that change litmus paper from blue to red</p>

3.4. Research procedures

Research procedure is explained to students, and a practice session was held at the start of the second week of the semester. Actual quiz games were administered every week from the second week of the semester until the last, with the exception of the midterm weeks and the final week. We have ensured there is no overlap between the 90 questions (10 weeks x 9 questions each week) on CQGs/PQs and the FECAT test, and also, that neither CQGs/PQs nor FECAT overlap with weekly homework or other assessment instruments. FECAT was administered in the very first and the very last week of the semester, and the Hake gain

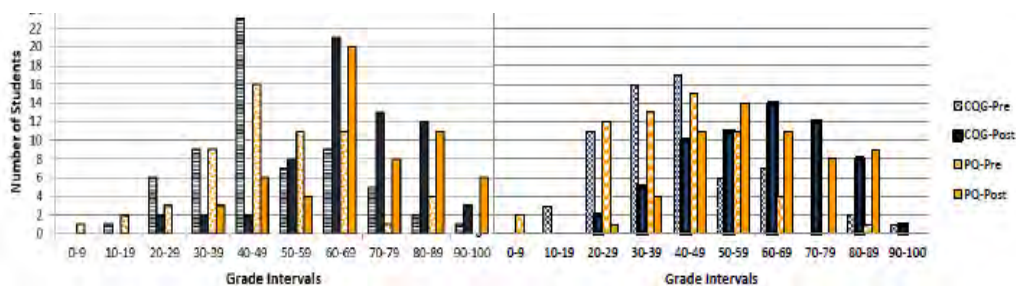
$$\text{Hake's Gain, } \langle g \rangle = \frac{\langle \text{Posttest Score} \rangle - \langle \text{Pretest Score} \rangle}{100 - \langle \text{Pretest Score} \rangle}$$

was computed (Hake, 1998, 2006, 2008; Hitt *et al.*, 2014; McKagan *et al.*, 2016) to calculate the learning gain. In addition to the overall Hake gain, which we have pursued as one of the research questions of interest from the outset of the study, we have computed the Hake gain for HOT and LOT set of questions.

4. Results

Figures 3(a) and 3(b) show the histogram of FECAT by comparing between pre-posttest for TG and CG of students' thinking skills as indicated by the types of questions (based on Bloom taxonomy) in FECAT. We found that the average performance for TG shifted from about 40% to 70%, slightly larger compared to CG, shifts from about 40% to 60% for students' HOT skills. In terms of students' LOT skills, the average performance of about 40% in pre-test to 70% in post-test was visible in both groups (Figure 5.9) with some quantitative differences between the two groups.

Figure 3. Comparison between pre-posttest for TG and CG of students' HOT (L) and (LOT (R) skills as indicated by the types of questions in FECAT; the number of students within the performance percentile.



We adopted Hake's gain expression (Hake, 1998, 2006, 2008; Hitt *et al.*, 2014; McKagan *et al.*, 2016) to assess the student learning gain. We found that both groups had positive Hake's gain, likely because of a similar motivational factor, as shown in Table 1 below. In a separate semester, we have conducted the Hake gain determination without administering any computer quiz games or paper quizzes, which we consider a true control group. The Hake's percent gain in that semester is substantially smaller for both groups, written here as percentage in (HOT, LOT) format - (CQG (33.6, 36.3), PQ (31.5, 37.2) and Control (19.1, 22.7) The pretests score on FECAT in both semesters (with and without the application of CQG and PQ) were nearly identical, since no major changes in the profile of students' population occurred.

5. Discussions

The results indicate that the students increased in their performance percentile group between the beginning (pre-test) and end of the semester (post-test). Besides, we found that both TG and CG have positive close Hake's gain for HOT and LOT thinking skills. According to Hopson (2001) and Fatokun *et al.*, (2016), and McFarlane *et al.*, 2002, technology-enriched and/or game environment classrooms have a positive effect on student acquisition of higher-order thinking skills.

Students' HOT skills are enhanced by the learning activities in class through the in-depth thinking process (similar to the finding by Lateef *et al.*, 2016). Both CQG and PQ are learning activities applied in class for a whole semester, so, we suspect that this might be the reason students with PQ also show almost the same value of Hake's gain. Three factors: identical motivating factors, imbalance of make-up of groups and high standard deviation might be the reason for the better result in CG.

6. Conclusion

Supplementing CQGs as an active learning strategy during the teaching and learning process can help students to develop HOT skills such as applying, analyzing, synthesizing and evaluating. CQGs help to enhance students' higher-order thinking skills and problem-solving skills. This study indicates that both CQG and PQ help to improve students' level of thinking skills and provide an alternative way of supplementing the traditional teaching over the true control group where neither CQG nor PQ is implemented at all. The implementation of CQG together with traditional lecture is suitable and effective for the university freshman students, particularly for students with education and culturally diverse background in UAE as well as students in other countries that have a similar background. We hope that this study can be a guidance to many other chemistry instructors and education authorities, particularly those dealing with the students who have a similar background.

References

- Adesoji, F. A. & Babatunde, A. G. 2008. *Investigating gender difficulties and misconceptions in inorganic chemistry at the senior secondary level*. International Journal of African & African American Studies 7(1): 1-7.

- Balawi, S., K. Khalaf, G. W. Hitt, *Leveraging Pedagogical Innovations for science, Technology, Engineering and Mathematics Education in the Middle East Context*, Advances in Engineering Education in the Middle East and North Africa 55, (2015), ed. by M. Abdulwahed *et al.*, ISBN 978-3-319-15323-0.
- Campbell, D. T., & Stanley, J. C. 1963. *Experimental and Quasi-experimental Designs for Research*. Chicago: Rand McNally College Publishing Company.
- Donald, J. G. 2002. *Learning to Think: Disciplinary Perspectives*. The Jossey-Bass Higher and Adult Education Series. San Francisco, CA: Jossey-Bass.
- Duron, R., Limbach, B. & Waugh, W. 2006. *Critical thinking framework for any discipline*. International Journal of Teaching and Learning in Higher Education 17(2): 160-166.
- Fatokun, K.V.F., Egya S. O. & Uzoechi B. C. 2016. *Effect of game instructional approach on chemistry students' achievement and retention in periodicity*. Eur. J. Res. Refl. Educ. Sciences 4(7).
- Felicia, P. 2011. *What is the relationship between digital games for learning and students' achievement?* EUN Partnership AISBL. December 2016
- Gokhale, A. A. 1995. *Collaborate learning enhances critical thinking*. Journal of Technology Education 7(1). <https://scholar.lib.vt.edu/ejournals/JTE/v7n1/gokhale.jte-v7n1.html> [25 April 2016]
- Gribbons, B. & Herman, J. 1997. *True and quasi-experimental designs*. Practical Assessment, Research & Evaluation 5(14):1-3. <http://PAREonline.net/getvn.asp?v=5&n=14> [29 July 2015]
- Hake, R. R. 1998. *Interactive-engagement versus traditional methods: a six-thousand-student survey of mechanics test data for introductory physics courses*. American Association of Physics Teachers 66(1): 64-74.
- Hake, R. R. 2006. *Possible Palliatives for the paralyzing pre/post paranoia that plagues some PEP's*. Journal of Multidisciplinary Evaluation 3(6): 59-71.
- Hake, R. R. 2008. *Can distance and classroom learning be increased?* International Journal for the Scholarship of Teaching and Learning 2(1), 21.
- Hitt, G. W., Isakovic, A. F., Fawwaz, O., Bawa'aneh, M. S., El-Kork, N., Makkiyil, S. & I. A. Qattan, I. A. 2014. *Secondary implementation of interactive engagement teaching techniques: Choices and challenges in a Gulf Arab context*. Physical Review Special Topics-Physics Education Research 10(2).
- Hopson, M. H.; Simms, R. L. & Knezek, G. A. 2001. *Using a Technology-Enriched Environment to Improve Higher-order Thinking Skills*. J. Res. Techn. Educ. 34(2): 109-119.
- Kapp, K. M. 2012. *The Gamification of Learning and Instruction: Case-Based Methods and Strategies for Training and Education*. New York: Pfeiffer: An Imprint of John Wiley & Sons.
- Lateef, A., Dhahr, M. A. & Latif. K. 2016. *Impact of higher order thinking skills of university students on their academic performance*. Science International (Lahore) 28(2): 2031-2035.
- Ling, S. S., Isakovic, A. F., Osman, K. Bt. & Halim, L. Bt. 2017. *The Uses of Instructional Strategies to Improve Students' Attitude to Learn Chemistry*. Journal of Educational Research, volume 37. [ISSN:0126-5261]
- Ling, S. S. 2018. *The Effectiveness of Computer-based Quiz Games Integrated with Lecture Method on Students' Achievement and Attitudes toward Introductory Engineering Chemistry*. PhD thesis. National University of Malaysia, Malaysia.
- McDonald, P., Straker, H., Schlumpf, K. & Plack, M. 2014. *Learning Partnership: Students and Faculty Learning Together to Facilitate Reflection and Higher Order Thinking in a Blended Course*. Online Learning: Official Journal of the Online Learning Consortium 18(4).
- McFarlane, A., Sparrowhawk, A. & Heald, Y. 2002. *Report on the Educational Use of Games*. London: Department for Education and Skills.
- McKagan, S., Sayre, E. & Madsen, A. 2016. *Normalized gain: What is it and when and how should I use it?* PhysPort, AAPT.
- Ramos, J. L. S., Dolipas, B. B., Villamor, B. B. 2013. *Higher order thinking skills and academic performance in physics of college students: A regression analysis*. International Journal of Innovative Interdisciplinary Research, issue 4.
- Rodrigues, A & Oliveira, M. 2008. *The Role of Critical Thinking in Physics Learning*. <https://core.ac.uk/download/pdf/38301359.pdf> [2 March 2019]
- Scott, S. 2009. *Perceptions of students' learning critical thinking through debate in a technology classroom: a case study*. Journal of Technology Studies 34(1): 39-44.
- Tiruneh, D.T., De Cock, M., Weldeslassie, A.G., Elen, J. & Jansen, R. 2017. *Measuring Critical Thinking in Physics: Development and Validation of a Critical Thinking Test in Electricity and Magnetism*. International Journal of Science and Math Education 15(4): 663-682.
- Trochim, W. M. K. 2001. *The Research Methods Knowledge Base*, 2nd ed. Cornell University Press.
- Van Der Werff, J. A. 2016. *Exploratory study of graduate-level instructor's perception of teaching critical thinking*. PhD dissertation. Kansas State University, Manhattan, KS.