Evaluating the Impact of Puzzle-Based Learning on Creativity, Engagement and Intuition of STEM Tertiary Students

Sergiy Klymchuk[a], Tanya Evans[b], Priscilla Murphy[a]. Julia Novak[b], Jason Stephens[b], Mike Thomas[b]

Auckland University of Technology[a], University of Auckland[b]

Corresponding Author Email: sergiy.klymchuk@aut.ac.nz

CONTEXT

In recent years, some universities in Australia, Europe and the USA, have introduced formal academic courses or seminars for their first-year STEM students based on a Puzzle-Based Learning (PzBL) pedagogical strategy introduced in Michalewicz and Michalewicz (2008). The feedback reported in several pilot studies was positive and primarily based on students' and lecturers' attitudes and perceptions. It indicated that PzBL could potentially increase students' engagement and enhance their creativity, general problem-solving skills and lateral thinking "outside the box". This paper attempts to evaluate those claims by rigorously measuring the impact of the PzBL.

PURPOSE OR GOAL

To select the best from the best, many high-tech companies often use puzzles at their job interviews. They believe that the ability to solve puzzles relates to the creative thinking needed for solving innovative real life problems. Our study investigated the impact on creativity, engagement and intuition of STEM students of a relatively short (one semester), but regular (weekly), PzBL intervention in lectures at four tertiary institutions in New Zealand.

METHODOLOGY/METHODS

This study was conducted as part of a mixed-methods design research project over 2018-2020. Several different types of instruments were used to gather data: comprehensive student pre- and post-test questionnaires, a content validation survey for the questionnaires, student focus group interviews, a lecturer questionnaire, and class observations. All statistical analyses, including Chi-square tests, *t*-tests, Friedman tests, parametric tests and Repeated Measures ANOVAs, were conducted IBM SPSS version 23.

ACTUAL OUTCOMES

The main findings are as follows. Students' behavioural engagement was significantly greater during the intervention. Perceptions of the utility value of PzBL activity improved at the end of the semester for all students. There were no significant changes in students' convergent thinking (problem-solving), intuition or creativity (originality, fluency and elaboration traits of the divergent thinking) over the intervention. This was probably due to the relatively short timescale of the intervention (around one hour in total over one semester).

CONCLUSIONS/RECOMMENDATIONS/SUMMARY

The results of this study suggest that with a relatively small effort lecturers can improve STEM student engagement by spending just a few minutes per lecture on PzBL activity during lectures—something that can be easily implemented, even for those who primarily teach using a traditional 'transmission' style. Moreover, such an intervention is highly amenable to scaling up with a relatively small development investment.

KEYWORDS

Engagement, Puzzle-based learning, STEM.

Introduction

In recent years, some universities in Australia, Europe and the USA have introduced formal academic courses or seminars for their first-year STEM students based on a Puzzle-Based Learning (PzBL) pedagogical strategy introduced in Michalewicz and Michalewicz (2008). Positive feedback from students, lecturers, and researchers on the regular use of pedagogical strategies employing non-routine problems and puzzles has been reported from several studies (see, for example, Falkner, Sherwood, & Michalewicz, 2012; Thomas, Badger, Ventura-Medina, & Sangwin, 2013; Klymchuk, 2017). In particular, the authors of the above studies noted that interesting non-routine problems, including puzzles, can engage students' emotions, creativity, and curiosity and also enhance their conceptual understanding, critical thinking skills, problem-solving strategies, and lateral "outside the box" thinking.

In this study, a non-routine problem means one for which the students do not have a ready-made method of solution they can apply, but the content is still within the scope of their knowledge base (Selden, Selden, Hauk, & Mason, 2000). These problems include puzzles that cannot be solved by rote, are presented in an entertaining way, look deceptively simple but have unexpected answers and counter-intuitive solutions, and hence are invaluable to promote reasoning that involves both logical and creative thinking (Thomas et al., 2013). They could also include paradoxes, which have a surprising, unexpected, counter-intuitive statement that looks invalid but in fact is true, and sophisms, which involve intentionally invalid reasoning that looks formally correct but contains a subtle mistake or flaw. Some illustrative examples of such non-routine problems/puzzles are:

• **The Race**. In a 100-metre race, athletes A, B, and C all run at a uniform pace. If A beats B by 10 metres and B beats C by 10 metres, by how much does A beat C? (Many students answer 20 metres, but the correct answer is 19 metres.)

• **The Average Speed**. You drive a car from A to B at a constant speed of 40 km/h. What should your constant speed be for the return trip from B to A if you want to obtain the average speed of 80 km/h for the whole trip? (Many students answer 120 km/h, but the correct answer is "infinite speed" or "impossible".)

There are clear links between skills developed by solving non-routine problems and puzzles and professional skills required in the workplace. Parhami (2008) argues that puzzling problems are plentiful in all research arenas regardless of discipline and, since many engineering problems are puzzle-like, engineering students should be exposed to them.

Student employability is another important aspect that relates to non-routine problem-solving. In addition to specific subject-based knowledge, many high-tech companies require good generic problem-solving and thinking skills from their employees. To select the best, companies often use puzzles at their job interviews. Management at high-tech companies believe that the ability to solve puzzles relates to the creative thinking needed for solving real-life problems in innovative ways. A classic example is Microsoft, where the goal of interviews is to assess general problem-solving ability rather than specific knowledge. Microsoft sees parallels between the reasoning used to solve puzzles and the thought processes involved in solving real-world innovative problems (Poundstone, 2000).

This paper investigates the impact on creativity, engagement and intuition of STEM students of a relatively short (one semester), but regular (weekly), PzBL intervention in lectures at four tertiary institutions in New Zealand.

Research Questions

The primary aim of the study was the evaluation of a strategic and innovative pedagogical intervention based on Puzzle-Based Learning (PzBL) in undergraduate STEM courses.

To fulfil this aim, we attempted to answer the following three research questions:

• Does the integration of non-routine problem-solving in lectures affect participants' engagement in lectures, and/or their ability to inhibit intuitive thinking and exhibit creative thinking?

• Are any observed effects moderated by individual differences such as demographic characteristics or prior ability?

• How do students react to the integration of non-routine problem-solving in their lectures?

Methodology and Methods

This research was conducted as a mixed-methods multi-study project over 2018-2020. Several different types of instruments were used to gather data: student pre- and post-test questionnaires, student focus group interviews, lecturer questionnaires, and naturalistic class observations.

In this study, there were 683 eligible undergraduate Year 1 and Year 2 students studying 12 STEM courses in astronomy, computing, engineering, and mathematics at four tertiary institutions in New Zealand: Auckland University of Technology (152 students), University of Auckland (440 students), Manukau Institute of Technology (25 students), and Whitireia New Zealand Polytechnic (66 students). During the first week of the first semester 2018, all students were invited to participate in the study via a verbal announcement in a lecture and an online invitation. Out of the 683 students, 100 students (approximately 15% of those invited) volunteered but only 64 students completed both pre-test and post-test questionnaires (a 9% response rate). Thus, the final sample included 36 males (56%), 27 females (42%), and one gender diverse (2%); 58 were younger than 25 years old (91%) and the remaining six participants were over 25 years old (9%). In addition to this sample of participants who completed both questionnaires, there were two smaller samples of participants: 12 student participants who volunteered for a one-hour task-based interview and nine lecturers involved in the intervention who completed the lecturer questionnaire by email. Participating students were rewarded with gift vouchers for their time. The project was approved by the University of Auckland Human Participants Ethics Committee on 05/06/18 (ref. number 021387) with subsequent approval by the other three tertiary institutions.

Students were exposed to the regular use of puzzles in lectures as a co-curricular activity over a 12-week semester. The lecturer participants were invited to the study and were given instructions and a selection of puzzles. They agreed to incorporate the puzzles activities in their lectures. Typically, 3–4 puzzles were presented to the students in one or two lectures every week. Students had a choice to solve puzzles either individually or in groups. Then a student or a lecturer showed a solution with a short explanation/discussion. The duration of this activity was 5–7 minutes a week. Hence, while the intervention involved solving and discussing a total of up to 40 puzzles over a 12-week semester, the intervention was short, comprising less than 5% of weekly lecture time. None of the other components of the course delivery (tutorials, assignments, tests, and examinations) were involved, as the intervention was designed by the project team to be easily incorporated into existing courses by slightly tuning the pedagogical practice in lectures.

The student pre-test (in week 1) and post-test (in week 12) questionnaires were designed to measure several psychological constructs and capacities (e.g., engagement, response inhibition, creativity). The questionnaires comprised 10 tasks/puzzles in each pre- and post-test and 30 questions in the pre- and 42 questions in the post-test questionnaires. We expected that students would take about 30 minutes to complete each questionnaire. The questions were from the following sources:

a) Engagement—two instruments were used: one based on the Expectancy-Value Theory (Eccles et al., 1983), which comprised 17 self-assessment items, assessing 3 latent factors related to solving non-routine problems: participants' self-efficacy, intrinsic interest, and utility value. The second questionnaire was based on Ahlfeldt et al. (2005) and was developed based on the USA's National Survey of Student Engagement (NSSE) (2000). It comprised 14 items, capturing a metric indicator of a 'global' measure of student engagement with the course.

b) Inhibition of intuition—an extended and slightly modified Cognitive Reflection Test (CRT) that included the original three questions from CRT (Frederick, 2005) as well as two questions inspired by Babai et al. (2015). An illustrative example from the Cognitive Reflection Task section of the pre-test questionnaire is below:

Question 1. If it takes 8 machines 4 minutes to make 8 widgets, how long would it take 100 machines to make 100 widgets?

c) Convergent thinking—the Convergence Task based on four counter-intuitive puzzles from Klymchuk (2001). An illustrative example from the Convergence Task section of the pre-test questionnaire is below:

Question 1. There are 10 New Zealand and 10 Australian coins of similar size and weight in a box. You take the coins out of the box without looking at them. What is the smallest number of coins you need to take out of the box to be absolutely sure that you have 5 coins of the same country, either New Zealand or Australian?

Questions/puzzles in the Cognitive Reflection Task and Convergence Task sections of the post-test questionnaire had very similar nature as in the pre-test questionnaire, some were slightly modified by changing the numbers and/or context. A Content Validation Survey completed by a group of seven experts established the equivalence of the questions/puzzles in the pre-test and post-test questionnaires.

d) Divergent thinking/creativity was assessed using an Alternative Uses Test from the well-established Torrance Test of creativity (Torrance, 1963) based on Guilford's model of creativity (1959). Specifically, participants were instructed to "spend 3 to 5 minutes" to "list as many possible uses for a Brick" or a "Shoe" in pre- and post-test. Responses were scored for two types of divergent thinking: 1) fluency (total number of the uses listed by each participant) and 2) elaboration (amount of the descriptive detail in participants' responses).

Given the importance of student engagement in this study, the foregoing self-reporting was triangulated by lecturer feedback and lecture observations. The nine lecturers involved in the intervention provided feedback on the engagement of their students by email response to a questionnaire. In addition, observers were present in five of the lectures in three different tertiary institutions and recorded the time spent in class on non-routine problems, the number of problems introduced in each lecture, the quality of delivery, and the active engagement of the students. This active engagement was measured in two ways. One method was a five-point scale, indicating the observers' perception of engagement (from low to high); the other was a more in-depth observation, sampling and recording student off-task behaviour. In each lecture, the observer visually scanned a sample of approximately 30 students, noting in a chart any students demonstrating off-task behaviour. This observation was repeated every five minutes, with particular attention given to the timing of the delivery of the non-routine problems.

The SPSS 25 program was used for all statistical analyses, which included Repeated Measures ANOVAs, Friedman test and regression analysis. The majority of the 64 students answered all of the questions but occasionally there were missing data.

Actual Outcomes

It is not possible to present here the actual data and analysis due to the page limit. We summarise here the main findings of the study that answer the research questions and refer to two recent journal articles that provide detailed analysis of some aspects of the study.

Engagement

• Students' behavioural engagement was significantly greater during the intervention. The evidence showed that they found the problems more engaging than the lecture itself, with fewer instances of off-task behaviour observed.

• The group with C grades in prerequisite courses appeared to be the most engaged with non-routine problem-solving.

NB. A comprehensive analysis of the impact of the PzBL intervention on student engagement in this study is presented in (Evans et al., 2020).

Self-efficacy, Intrinsic interest and Utility value

• Students mostly indicated positive levels of self-efficacy in solving non-routine problems. They saw themselves as capable of solving non-routine problems, perceived a value and enjoyed doing so, indicating a positive emotional disposition.

Intuition

• Even though students saw the importance of inhibiting intuitive thinking, it did not change significantly over time.

Convergent thinking

• There were no significant changes in students' convergent thinking (problem-solving) ability over time.

Creativity

• There were no significant changes in students' creativity (originality, fluency, and elaboration traits of divergent thinking) over the intervention.

NB. A qualitative analysis of students' perceptions on creativity in the PzBL intervention in this study is presented in (Murphy et al., 2020).

Group differences

• Grades in prerequisite courses did not significantly influence over time student self-efficacy in, emotional disposition toward, or perceived value of non-routine problem-solving. However, students with prior B- and C-band grades reported a small increase in self-efficacy and intrinsic interest.

• The results suggest that the intervention may have been more effective for males than females and this hypothesis could be investigated in future research.

Student perceptions about learning

• Students agreed that solving non-routine problems was useful for their learning and was able to enhance their creative and innovative thinking abilities. They talked about the need to "think outside the box" and think holistically rather than be focused on a single approach.

• Perceptions of the utility value of non-routine problem-solving improved at the end of the semester for all students.

• The students strongly agreed solving non-routine problems in their courses would be beneficial to their future learning, as well as their careers and other areas of life.

We anticipated a high level of student engagement during the PzBL intervention and this was realised. The feedback by the vast majority of the students also confirmed that they saw valuable benefits from the PzBL activities for their future employment. We also expected that the PzBL intervention would enhance students' creativity and ability to inhibit their intuition. However, the increase in students' creativity and ability to control their intuition before and after the intervention was not statistically significant. One possible reason for this could be the small timescale of the intervention with simply not enough time to exhibit such a change. Although the intervention was regular, each week for three months, the total time was only about an hour. Having spent 10+ years in formal education where there was little or no attention paid to creativity in STEM subjects, perhaps students need more time to unlock it. Many researchers point out that (primary) intuition is strongly present and very resistant to change (e.g., Fischbein, 1982; Thomas, 2015). Hence, these important aspects of enhancing creativity and the ability to inhibit intuition when necessary require further attention by researchers.

Conclusions and Recommendations for Practice

Increasing behavioural and cognitive engagement of students in tertiary institutions is extremely important, and behavioural and cognitive engagement are often used as indicators of improved learning (Kahu & Nelson, 2018; Watt & Goos, 2017). A primary outcome of this research is that an intervention employing non-routine problems significantly improved student engagement and induced positive attitudes to the solution process. Hence, we would recommend that tertiary institutions seriously consider initiating a similar pedagogical strategy in their STEM subjects. While a few universities may already offer optional seminars or even compulsory courses based on PzBL for their first-year STEM students, this is far from the norm. University lecturers may think it is too difficult to adopt an approach to teaching that fosters student engagement, given their limited resources, large class sizes and the significant amount of time and effort required for class preparation. However, the results of this study suggest that with a relatively small effort, lecturers can improve student engagement by spending just a few minutes per lecture on non-routine problem-solving during STEM lectures—something that can be easily implemented, even for those who primarily teach using a traditional “transmission” style. Moreover, such an intervention is highly amenable to scaling up, with a relatively small development investment and easy transfer to other tertiary STEM courses since it requires minimal time allocation or lecturer preparation. The student voice in the process is also very important, and regular student feedback—both formal and informal—will make the intervention more meaningful and successful.

In order to benefit fully, we recommend that lecturers consider working in small groups when implementing this practice. The research literature shows the value of small communities of inquiry as part of professional development (Jaworski, 2006) in order to discuss ideas and provide support (Barton et al., 2014) as lecturers engage in critical inquiry into their practice. This would also assist with planning, including preparation of the non-routine problems that would be relevant, allowing individuals to provide feedback to the group, and possibly enabling peer observation to guide improvement (Jaworski, 2006). There are numerous resources on popular mathematics and science from which a lecturer can choose puzzles for a specific course: websites, books, journals, magazines.

In conclusion, we can say that the students here were confident that the experience of solving non-routine problems in lectures would be of value to them, both for their current learning and for their future education and careers. Furthermore, they enjoyed the experience and had a positive emotional response to it, which made them feel more capable of solving non-routine problems. That the greatest effect on both self-efficacy and intrinsic interest appeared to be for students with lower prior achievement should add to the appeal of the approach for many institutions because it has the potential to increase the student retention rate. Overall, the students were emotionally engaged and interested rather than disengaged, making them far less likely to drop out of courses (Blondal & Adalbjarnardottir, 2012). All of these positive factors contribute to the strength of the recommendation that other educators give the introduction of this kind of intervention a serious consideration.

Limitations

One of the main limitations of this research was that it was a reasonably short timescale intervention. Neither more time could be spent on it, nor the PzBL could be integrated deeper in the courses, since the puzzles were outside of the standard content of those courses. There was also a relatively low response rate (leading to an overall small sample size) despite a very positive attitude of the vast majority of students towards the PzBL intervention and their high level of engagement. Possible reasons for this include the requirement to attend most lectures to be eligible for the study (typical attendance was around 50%), the comprehensive nature of pre- and post-test questionnaires, and the challenging tasks that some students were not confident to do. Due to a small sample size, we sometimes were not able to make enough observations to detect significant differences, for example in the parametric test. Another limitation was the use of a non-probability sampling—a convenience sampling method with students self-selected for the study was used instead of a random sampling.

References

Ahlfeldt, S., Mehta, S., & Sellnow, T. (2005). Measurement and analysis of student engagement in university classes where varying levels of PBL methods of instruction are in use. *Higher Education Research & Development, 24*(1), 5–20.

Babai, R., Shalev, E., & Stavy, R. (2015). A warning intervention improves students' ability to overcome intuitive interference. *ZDM Mathematics Education*, *47*(5), 735–745.

Barton, B., Oates, G., Paterson, J., & Thomas, M. O. J. (2014). A marriage of continuance: Professional development for mathematics lecturers. *Mathematics Education Research Journal, 27*(2), 147–164.

Blondal, K. S., & Adalbjarnardottir, S. (2012). Student disengagement in relation to expected and unexpected educational pathways. *Scandinavian Journal of Educational Research*, *56*(1), 85–100.

Eccles, J., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J., & Midgley, C. (1983). Expectancies, values and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives* (pp. 75–146). W. H. Freeman.

Evans, T., Klymchuk, S., Novak, J., Murphy, P., Stephens, J. M. & Thomas, M. (2020). Engagement of undergraduate STEM students: The influence of non-routine problems. *Higher Education Research & Development*. Published online 30 October 2020. https://doi.org/10.1080/07294360.2020.1835838

Falkner, N., Sherwood, R., & Michalewicz, Z. (2012). Teaching puzzle-based learning: Development of basic concepts. *Teaching Mathematics and Computer Science, 10*(1), 183–204.

Fischbein, E. (1982). Intuition and proof. *For the Learning of Mathematics*, *3*(2), 9–24.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*(4), 25–42.

Guilford, J. P. (1959). Traits of creativity. In H. H. Anderson (Ed.), *Creativity and its cultivation* (pp. 142–161). Harper & Brothers.

Jaworski, B. (2006). Theory and practice in mathematics teaching development: Critical inquiry as a mode of learning in teaching. *Journal of Mathematics Teacher Education, 9*(2), 187–211.

Kahu, E., & Nelson, K. (2018). Student engagement in the educational interface: Understanding the mechanisms of student success. *Higher Education Research & Development, 37*(1), 58–71.

Klymchuk, S. (2001). *Money puzzles: On critical thinking and financial literacy*. Maths Press.

Klymchuk, S. (2017). Puzzle-based learning in engineering mathematics: Students' attitudes. *International Journal of Mathematical Education in Science and Technology*, *48*(7), 1106–1119.

Michalewicz, Z., & Michalewicz, M. (2008). *Puzzle-based learning: An introduction to critical thinking, mathematics, and problem solving.* Hybrid Publishers.

Murphy, P., Evans, T., Klymchuk, S., Novak, J., Stephens, J., Thomas, M. (2020). University STEM students' perceptions of creativity in non-routine problem-solving. *Australian and New Zealand Industrial and Applied Mathematics Journal (ANZIAM Journal)*, 61 (EMAC2019), C152-C165.

National Survey of Student Engagement. (2000). *Improving the college experience: National benchmarks of effective educational practice.* Indiana University Center for Postsecondary Research and Planning.

Parhami, B. (2008). A puzzle-based seminar for computer engineering freshmen. *Computer Science Education*, *18*(4), 261–277.

Poundstone, W. (2000). *How would you move Mount Fuji? Microsoft's cult of the puzzle—How the world's smartest companies select the most creative thinkers*. Little Brown and Company.

Selden, A., Selden, J., Hauk, S., & Mason, A. (2000). Why can't calculus students access their knowledge to solve nonroutine problems? In E. Dubinsky, A. Schoenfeld, & J. Kaput (Eds.), *Research in collegiate mathematics education: IV. Issues in mathematical education* (Vol. 8, pp. 128–153). American Mathematical Society.

Thomas, C., Badger, M., Ventura-Medina, E., & Sangwin, C. (2013). Puzzle-based learning of mathematics in engineering, *Engineering Education*, *8*(1), 122–134.

Thomas, M. O. J. (2015). Inhibiting intuitive thinking in mathematics education. *ZDM Mathematics Education*, *47*(5), 865–876.

Torrance, E. (1963). *Education and the creative potential*. University of Minnesota Press.

Watt, H. M. G., & Goos, M. (2017). Theoretical foundations of engagement in mathematics. *Mathematics Education Research Journal, 29,* 133–142.

Acknowledgments

This research was supported by a Teaching and Learning Research Initiative grant from the New Zealand Council for Educational Research.

Copyright statement