EXPLORING LOG DATA FOR BEHAVIOUR AND SOLUTION PATTERN ANALYSES IN A SERIOUS GAME

Chioma Udeozor[a], Fernando Russo Abegão[a], Jarka Glassey[a*]

[a] *Merz Court, School of Engineering, Newcastle University, NE17RU. Newcastle upon Tyne, United Kingdom.*

*jarka.glassey@newcastle.ac.uk

EXPLORING LOG DATA FOR BEHAVIOUR AND SOLUTION PATTERN ANALYSES IN A SERIOUS GAME

Abstract

Digital games are interactive, which makes them highly engaging for players. The adoption and use of digital games in higher education are on the rise with many researchers and educators developing and deploying these in classrooms. As a relatively new pedagogical tool, some aspects of the use of games for learning such as measurement and assessment of learning are still under research. Although assessment of performance and learning in digital games are commonly done with pre and post-game tests, interest in growing in the use of gameplay log data as an alternative and valid means of measuring the performance of students in digital games. A few studies have utilized log data to measure the performance of students in general knowledge and skills but limited studies exist where game log data were used to measure domain-specific competencies. This empirical study describes the use of game log data for measuring the behaviours and performance of engineering students in the Cosmiclean game, a serious game designed to teach the principles of separation and recycling operations. Using the data from first year engineering students from two European institutions, sequential behaviour pattern analysis and performance assessment of students solutions in the game are presented. The findings of this study highlight the behaviours and gameplay strategies of students in the game environment, and these would be particularly useful to game designers, educators and researchers in the field of game-based learning.

Keywords: Serious game, log data, performance, engineering, recycling, separation operations.

1 INTRODUCTION

1.1 Digital Game-Based Learning

The use of digital games for education is gaining considerable interest as a pedagogical tool that is engaging and interactive. The adoption and use of digital games in higher education are also on the rise with many researchers and educators developing and deploying these in classrooms. The COVID-19 pandemic is also intensifying the trends towards serious games use. With the increase in remote learning and limited access to school laboratories and workshops, some institutions have explored the use of gamified virtual laboratories (Glassey & Magalhães, 2020). Students are now provided access to virtual laboratory environments from home, hence carrying on with their education without violating social distancing measures.

The interest in game-based learning grew out of the belief that digital games have motivational and cognitive qualities that could be beneficial for education (Wouters & van Oostendorp, 2013), and these have led to the use of educational (serious) and entertainment games for pedagogical purposes. Digital games are interactive which makes them highly engaging for players. This quality makes them particularly impactful in education as they allow for problem-based and experiential learning opportunities that are of considerable importance for learning. However, as a relatively new pedagogical tool based on technology that is in continuous evolution, many aspects of the use of digital games for classroom use are still under research. While some studies have focused on understanding the perceptions of students towards digital game-based learning (Thanasi-Boçe, 2020; Udeozor, Russo Abegão, Gill, & Glassey, n.d.; Udeozor, Toyoda, Russo Abegão, Gill, & Glassey, submitted; Yue & Tze, 2015) others have explored the effectiveness of digital game-based learning (DGBL) as a pedagogical tool (Perini, Oliveira, Margoudi, & Taisch, 2018; Suescún et al., 2018; Xenos

& Velli, 2020). A few others, like the current work presented herein, are exploring ways of measuring learning performance with DGBL (Hou, 2012; Kang, Liu, & Qu, 2017; Loh, Sheng, & Li, 2015; Udeozor, Russo Abegão, et al., n.d.)

Assessment of learning is an integral part of what happens in school environments. In order to make DGBL more attractive to educators as well as students, it is important to determine valid and reliable means of measuring what students know and do in the game environment. Generally, when digital games are used for instructional purposes, assessment of learning can either take place *ex situ* (outside the game; external) or *in situ* (within the game; internal) (Loh, Sheng, & Ifenthaler, 2015). External assessment involves the use of items such as quizzes, tests, essays, peer assessment, which are separate from the game, to determine the learning performance of students after playing a learning game. These assessment types have been found to be the most commonly used in game-related studies due to their ease of use (Bellotti, Kapralos, Lee, Moreno-Ger, & Berta, 2013; Loh, Sheng, & Ifenthaler, 2015). Nonetheless, the use of external assessment in DGBL has been criticised as being isolated from the learning context. It is believed that using these forms of assessments misses out on the opportunity for performance-based assessment that is afforded by the game, hence failing to measure more complex skills and competencies that are otherwise difficult to measure (Groff, 2018; Shute & Ke, 2012). As an alternative, in-game assessment methods such as game scoring (Bellotti et al., 2013; Moseley, 2013), log data analysis (Kerr & Chung, 2012; Loh, Sheng, & Li, 2015; Westera, Nadolski, & Hummel, 2014), and integrated or stealth assessment (Almond, 2015; Kim, Almond, & Shute, 2016; Shute, Wang, Greiff, Zhao, & Moore, 2016) have been proposed for game-based learning. Therefore, this study explores the use of game log data to determine solution patterns used by students in the game, in order to investigate the possibility of assessing problem-solving skills of students in a serious game using a performance-based method.

1.2 Performance-based assessment with game log data

A few studies in the field of game-based learning have explored the use of game log data for measuring the learning performance of students in a serious game. Westera et al. (2014) used game log files of students to identify different gaming behaviours. In their study, regression analysis was also used to determine the relationship between the gaming behaviours and post-test scores of students. They found that video access rates and overall activity rates of students in the game were good predictors of learning efficiency. Game log data have also been used to predict expert-novice performance in serious games (Loh & Sheng, 2015; Loh, Sheng, & Li, 2015). In their studies, Loh and colleagues mapped the action sequences performed by university students in serious games. By calculating the Jaccard coefficients of the in-game actions, they were able to group students into expert and novice categories. In other studies, cluster analysis was used to measure and identify key features of the performance of primary school pupils in a math game (Kerr & Chung, 2012). Using the log data from an educational game called Save Patch, Kerr and Chung were able to determine the gameplay strategies of students. In the study of Hou (2012), cluster analysis was also used to identify the behavioural patterns of gameplay using the game log data of primary and high school students in a massively multiple online role-playing game (MMORPG). The study identified three clusters of gamers based on their gameplay activities: highest participation gamers, high-participation gamers, and ordinary-participation gamers. Another study (Smith, Hickmott, Southgate, Bille, & Stephens, 2016) used log data to visualize the decisions of players during gameplay. Log data were analysed to identify areas where university students spend the most time in the game environment, gameplay sequences used as well as their use of training videos.

Overall, game log data provide detailed action-based information of students in game environments. It has been successfully used to measure and identify different aspects of learning and gameplay performance. This successful use of log data for performance assessment has also been seen across different types of educational and entertainment games. In the present study, game log data will be used to explore the behaviours and performance of engineering students in a serious game called Cosmiclean game. A few studies have reported the use of game log data for assessment in professional academic disciplines. However, to the best of our knowledge, in engineering disciplines no study has been found to use gameplay log data for assessment of engineering skills and competencies. Hence, this study presents the use of unstructured log data from the gameplay of engineering students to measure their engineering knowledge and skills based on their performance in a recycling game. The findings of this study would be particularly useful to game designers, educators and researchers in the field of game-based learning as they provide a better understanding of gameplay behaviours, strategies and performance of students in a game-based learning environment. These should hence inspire better design approaches for learning games as well as highlight some considerations when using digital games for performance assessments.

1.3 Cosmiclean game

Cosmiclean game, also known as the Recycling game (https://recyclegame.eu/the-game/), is a serious game with strategy, puzzle, and adventure elements. It is a product of the European Institute of Innovation & Technology – Knowledge and Innovation Community (EIT-KIC Raw Materials) game project of the European Union Horizon 2020 (EUH2020) and was designed by LuGus Studio, Belgium. The game was designed to expose players to the science and challenges of waste separation and recycling using principles of automated industrial processes. With 56 levels of increasingly challenging gameplay, Cosmiclean game uses high-quality graphics and an engaging narrative to provide a fun game experience.

The game places the player in the position of an artificial intelligence (AI) trying to save a stranded damaged spaceship carrying a tone of waste across the galaxy. With no resource available to repair the ship, the AI resorts to using the waste cargo. The player has to sort the waste materials and use them to repair parts of the damaged ship as shown in Figure 1. With a working 3D printer on the ship, the player is also able to design and print different separators needed to sort the waste.

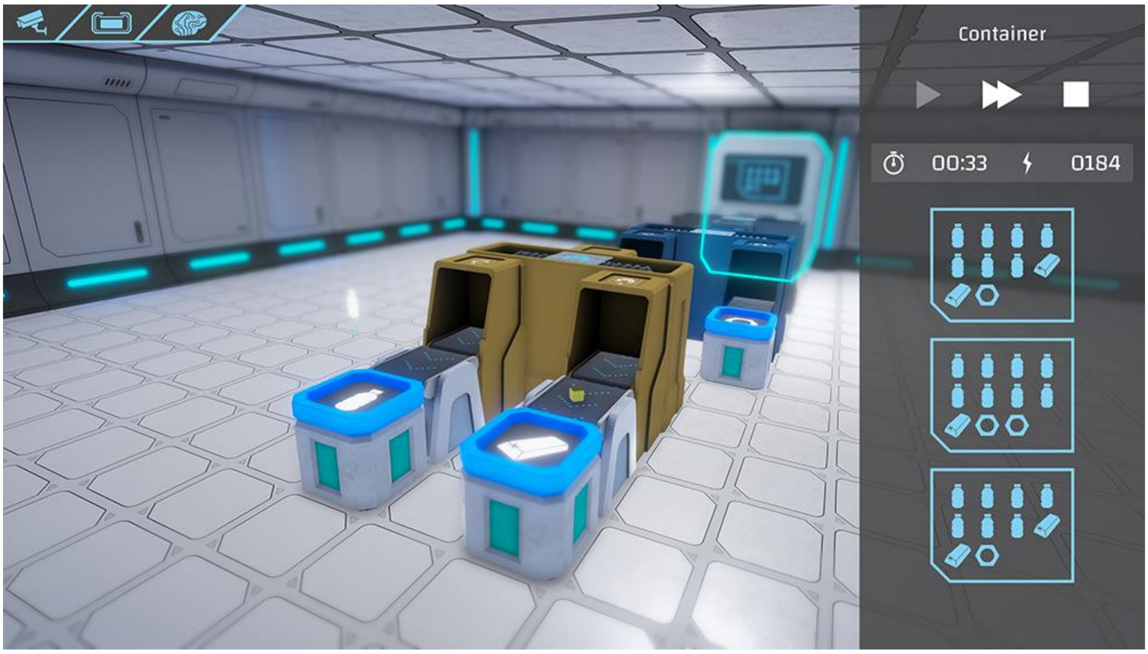

Figure 1: Screenshot of the separation process in the Cosmiclean game

When viewed from an engineering educational point, the Cosmiclean game teaches the heuristics of separation processes through waste recycling. The players have to determine and use relevant resources (including energy and processors) in the correct sequence to separate mixtures of waste by exploiting different properties of the materials at each level of the game. Although designed for the wider public, the gameplay tasks are closely aligned with some core modules of chemical engineering education, such as the principles of separation operations and elementary principles of project-based plant design. In the game, players make strategic decisions on which processors (i.e. equipment) to use, optimal sequencing, as well as the configuration of each processor to ensure the efficiency of the separation and the recovery of the raw materials from the waste cargos. This game was selected for this study because it was designed by engineering experts and has won the Comenius-EduMedia Siegel award 2019 for its outstanding content and pedagogical design (https://comenius-award.de/).

1.4 Aim of the study

This study aims to explore the use of game log data to assess the behaviour, performance and solution patterns of engineering students in the Cosmiclean game. The following research questions will be answered in this study:

- What trend exists in the sequential behavioural patterns of engineering students across the game?
- What performance clusters can be identified from the gameplay log data?
- What are the characteristics of the solution patterns obtained?

2 METHOD

2.1 Participants

A total of 58 chemical engineering students from KU Leuven and Imperial College London took part in this study. There were 39 male and 18 female students. 97% of them were between the ages of 20 and 29 years as shown in Table 1. Convenience sampling method was used to recruit these participants (Creswell, 2011). This cohort of students was selected because they are most likely to benefit from the gameplay, as the game serves as a practical introduction to the principles of separation operations, a core module taught in the 2nd year of study. They were also in the position to draw on their knowledge of sustainability while solving the game tasks.

Table 1: Participant demographic data

		Absolute Frequency
Gender	Male	39
	Female	18
	Unspecified	1
Age	Under 20 years old	2
	20-29 years old	56
Prior gameplay	Yes	54
	No	4
Gaming Habits per week	Less than 5 hours	25
	5-10 hours	21
	11-20 hours	7
	21-30 hours	4
	31 hours or more	1

2.2 Context

This study serves as a follow up on a previous study from the authors that looked at the relationship between gameplay performance and perceptions of students towards learning games (Udeozor, Russo Abegão, et al., n.d.). The Cosmiclean game was used for the study and data collection took place during the period. Participants were required to solve as many levels of the game as they could, but strive to solve a minimum of 25 levels within the 2-weeks timeframe given. They were asked to solve the tasks as a chemical engineer would, paying attention to the sequencing and configurations of the processors selected. Since the game does not provide feedback on the efficiency of solutions, participants would have to judge by themselves whether their solutions were optimal or not. Two factors were used to determine the solution patterns and the efficiency of the performance of students in the game. These were the average energy used for solving the tasks, and the number of resources used per level.

2.3 Data Collection

An online survey and game log files were used for data collection. To evaluate the game experiences of students, a questionnaire was used. The questionnaire collected socio-demographic data and game experiences as presented in Table 1. To measure the behaviour patterns in the game, gameplay log files were collected and analysed. The data collected contains students unique identifiers, all levels completed, energy expended and the resources used in each level. In this context, resources are viewed as every piece of equipment utilized during the process. This includes processors that transform materials (e.g. sieves, boilers and shredders) and non-processors that simply transport and hold materials (e.g. conveyors and receptors).

2.4 Data analysis and coding

To perform sequential pattern analysis of gameplay activities, log files were cleaned and processed in Rstudio 4.0 before analysing in SPSS. A descriptive analysis of the gameplay actions of the students is presented in Table 2.

Table 2: Descriptive results of gameplay activities of students.

	Mean	Std Dev	Maximum	Minimum
Levels Completed	25.9	14.6	57	3
Energy Expended per Level completed	53.5	32.4	202.8	16
Unique Solutions per Level	6.2	4.4	26	1

The game has 57 levels and a total of 11 resources (2 non-processors and 9 processors) available to use. To solve each level, participants were required to drag and drop suitable resources to the recycling line. They were to choose from a list of provided resources which differ from level to level. For each level solved by a student, the type and sequence of assembly of the resources were logged. To analyse these logged data, the string codes used were converted to numeric codes as shown in Table 3.

Table 3: Resources available in the game

Resources	Function	Symbol	Code
Conveyor	Connects two resources and transport materials from one to the other	A	1
Receptor	Collects predefined material	N	2
Sieve	Separates materials by size	J	3
Melter	Separates materials by melting temperature	M	4
Magnet	Separate ferrous metals from other materials	I	5
Shredder	Separates materials by size reduction	H	6
Eddy Current Separator	Separated non-ferrous metals from other materials	E	7
Stream Separator	Separates materials based on state of matter	L	8
Boiler	Separates materials by boiling temperature	G	9
Dissolver	Separates materials by solubility	D	10
Centrifuge	Separates materials by density	C	11

The resources in Table 3 are presented in order of appearance in the game. For instance, Level 1 of the game requires players to connect the material feed to the receptor as shown in Table 4. For this level, only the conveyor was needed and provided. For Level 5, players were provided sieves and conveyors and were required to separate 3 different materials given their particle sizes, whereas for Level 12, conveyors, sieves, melters and magnets were provided and players were required to choose and sequence the most suitable resources to separate glass, iron, and bricks of similar particle sizes.

Table 4: Examples of logged sequential solutions (coded).

Student number	Level 1	Level 5	Level 12
1	1-1-1-2	3-3-2-2-2	4-4-2-2-2
2	1-1-1-2	3-1-3-2-2-2	4-4-2-2-3
3	1-1-1-2	3-2-3-2-2	4-2-4-2-2
4	1-1-1-2	1-1-1-3-2-1-1-3-2-2	4-4-2-2-2
5	1-1-1-2	3-2-3-2-2	4-2-4-2-2

6	1-1-1-2	3-2-1-3-2-2	4-2-4-2-3
7	1-1-1-2	1-1-1-3-2-1-3-2-2	4-2-4-2-4

As shown in Table 4, different solutions patterns were possible at each level of the game. The solution to Level 1 was the same for all students, 1-1-1-2, which means students placed the resources in the order: conveyor-conveyor-conveyor-receptor. For the other two levels presented, there were multiple solutions by the students indicating the increase in task complexity and resources used.

3 RESULTS AND DISCUSSIONS

3.1 Gameplay Sequential Behaviour Pattern

To understand the sequential behaviour patterns of students across the entire level of the game, the coded sequential solutions for all levels were analysed, and all unique solutions were extracted. High variabilities were found in the number of unique sequential solutions across the entire 57 levels of the game as shown in Figure 2. Additionally, more unique solutions to individual levels were seen in the first few levels of the game compared to the later levels.

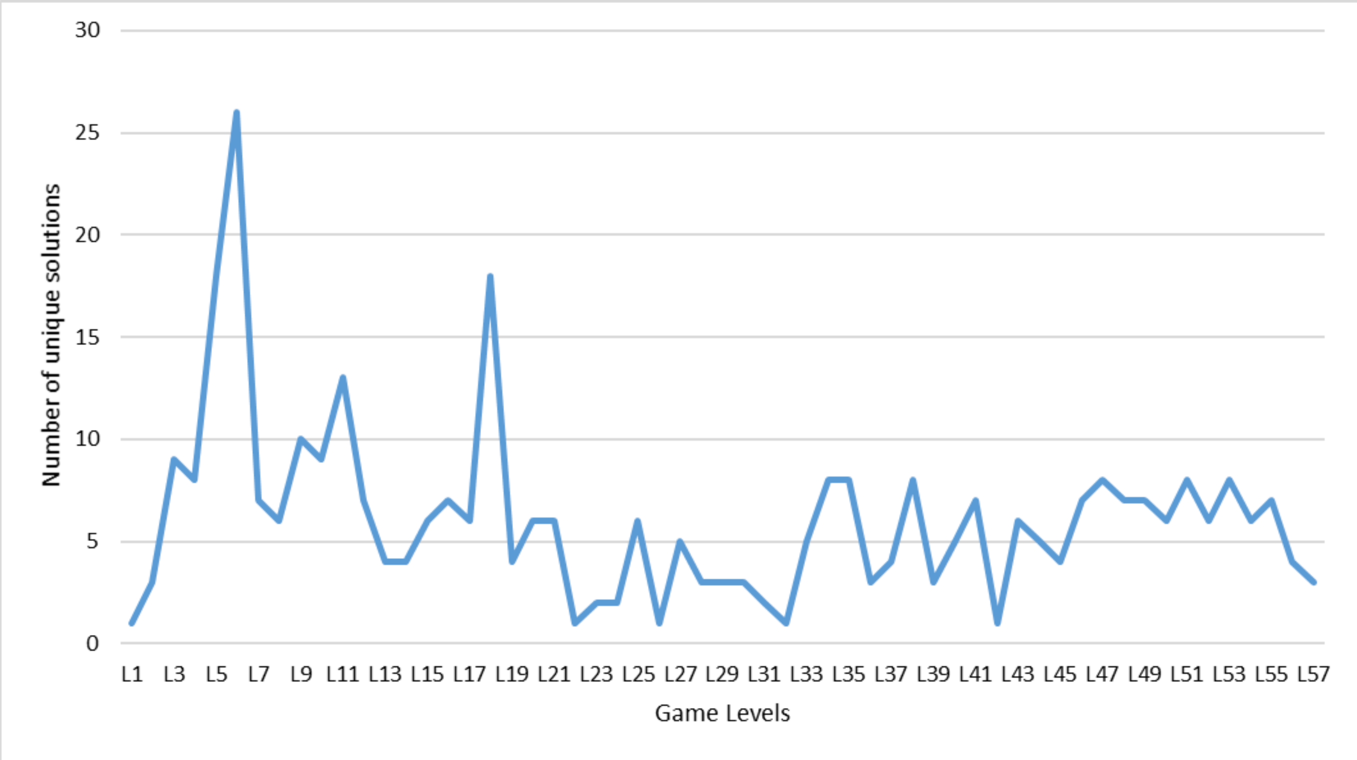

Figure 2: Sequential Solution Pattern across game levels

This implies that different students used different sequences of solutions and, in some cases, varying numbers of resources for solving a particular level. Given the ease of the tasks in the first levels, it was unexpected to observe the highest variability in the first levels of the game. However, this pattern can be explained by students engaging in explorative behaviours in the game environment as they familiarize themselves with the game in the early stages of the gameplay. This resonates with the findings of Kang et al. (2017) who identified several exploration activities of students during the first stages of their gameplay. Considering that students were not given any prior training on the game mechanics, it is logical that they explore how the game works and in the process, attempt different valid routes to solve the game tasks. As students progressed through the levels, the explorative patterns diminished despite the increased complexity of the tasks of those levels. This seems to indicate that students gradually learnt and evolved towards more optimal solutions during the game, and therefore, the cohorts progressed to a lower variability of strategies.

While these findings depict a real picture of the interactions of students with a learning game, the variability in solutions could be minimised by providing initial training time and pre-study gameplay. However, as the goal of the study was to understand the behaviour pattern of students in a serious game, pre-study training was not necessary as it would have biased the results obtained. This finding is particularly relevant to game designers, educators and researchers who work in the field of DGBL as it provides an understanding of what to expect from students when exposed to an unfamiliar learning game environment. It also points to the need to allow a considerable amount of gameplay time when using games for performance assessment. Measuring performance based on a few hours of gameplay might fail to provide valid and reliable insight into the competencies and skills limitations of students as they would most likely be at the exploratory phase of gameplay.

3.2 Clustering Performance

With the overall sequential behaviour pattern established, the next step was to assess the solutions and identify the different groups of solutions as well as the performance of the students. A 2-step cluster analysis involving hierarchical and K-means clustering were used to partition students into groups based on the performance on game tasks. Cluster analysis has been explored in recent studies as a means of assessing the performance of students from game log data (Kerr & Chung, 2012; Lin, Hsieh, Hou, & Wang, 2019). Given that, on average, students completed 25 levels, only the data of the students who completed the first 25 levels were selected in order to make the analysis more consistent.

A total of 32 students completed the first 25 levels of the game. To determine the optimal number of groups of students based on their gameplay performance, a hierarchical cluster analysis (Cohen, Manion, & Morrison, 2017) was first performed. The variables considered here to measure performance were the average number of resources used, and the average energy expended in solving the tasks. To begin the analysis, the scores on both variables were first standardized to make the relative weight of each variable equal. The Ward's Method with a dendrogram diagram was used for this (Cohen, Manion, & Morrison, 2017). From the dendrogram in Figure 3, the optimal number of clusters in the data was found to be three.

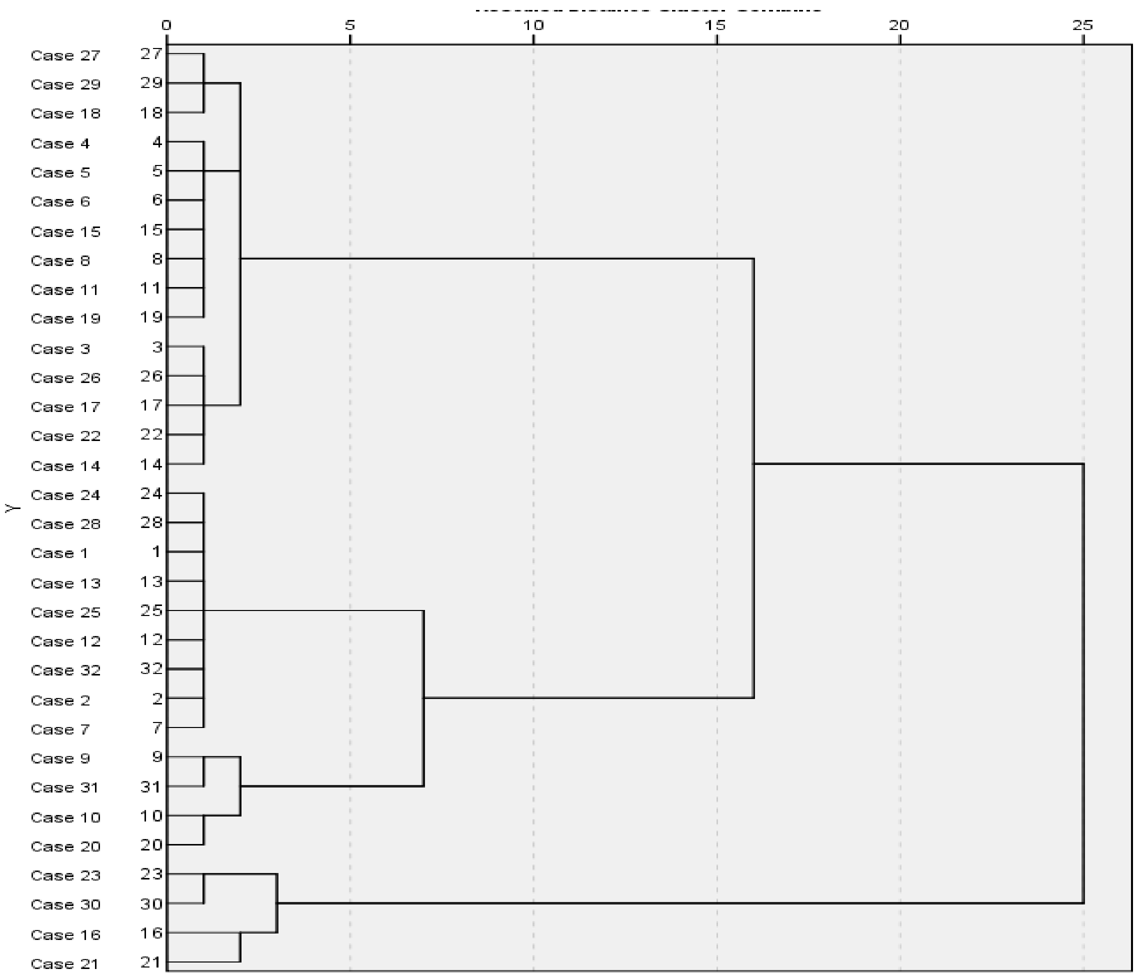

Figure 3: Dendrogram using Ward's Method

For the next stage, a 3-cluster K-means analysis was performed. From the results, it was found that Cluster 1 consists of students who expended less energy and the highest number of resources to solve all 25 levels. Cluster 2 on the other hand consist of students who used the least amount of resources and the highest amount of energy. The 3rd Cluster consists of students who used the least amount of energy and fewer resources in their solution. Cluster 1, 2 and 3 were labelled ‘Optimal Energy’, ‘Optimal Resources’, and ‘Optimal Solution’ groups respectively, as shown in the statistics presented in Table 5. To visualize the differences in performance of the three groups, the z-scores of the average energy and resources used per level were plotted and presented in Figure 4. With the values of average energy and resources given equal standards, i.e. mean of zero, this figure provides a visual of the performance of the groups relative to each other.

Table 5: Cluster Characteristics

	Cluster 1 Optimal Energy		**Cluster 2 Optimal Resources**		**Cluster 3 Optimal Solution**	
	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Average Energy	49.6	9.7	107.1	17.9	46	10
Average Resources	5.09	0.15	4.55	0.15	4.65	0.12
	N=5		N=4		N=23	

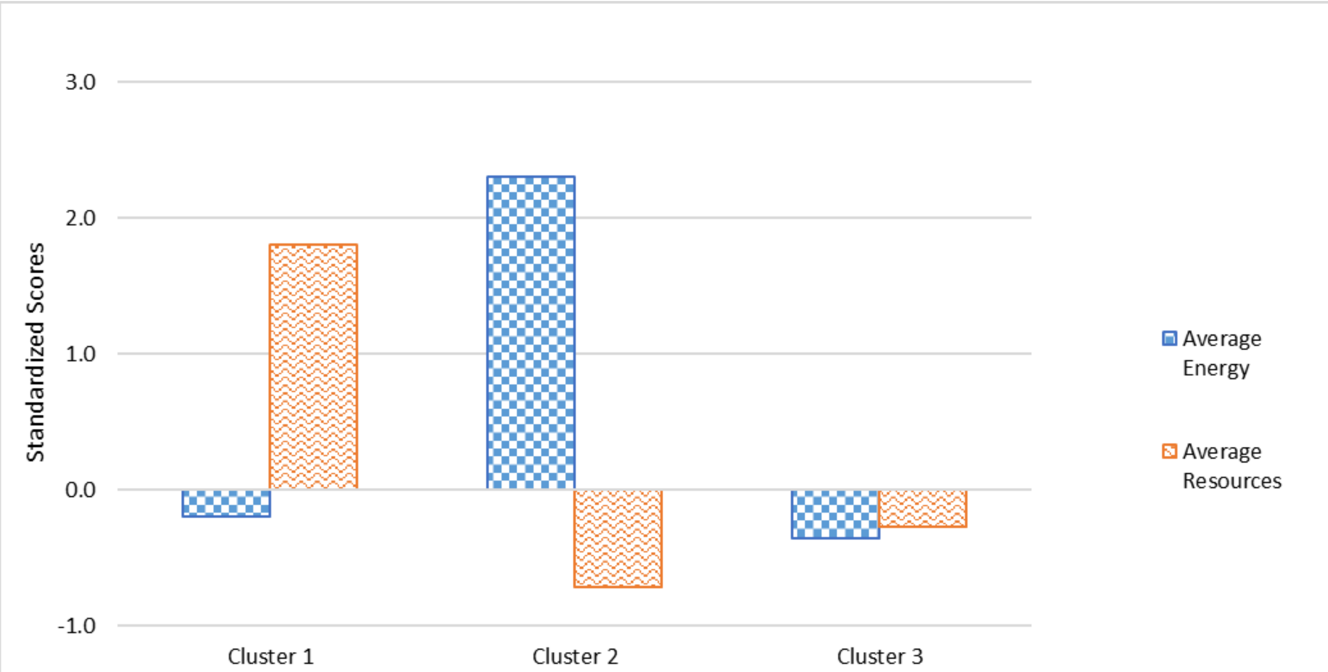

Figure 4: Visualization of cluster performance.

The Optimal Solution group performed considerably better than the other two groups. Their solutions are considered the most sustainable of the three, a factor that is of utmost importance in the chemical engineering domain (Glassey & Haile, 2012). The solutions of the students in this group involved the use of few resources, fewer than those of the Optimal Resource group but similar to those of the Optimal Energy group as shown in table 5. This implies that their solutions were thought through, consisting of mainly relevant resources and the use of optimal sequences. Additionally, the students in this group considered the energy requirements of the resources used in their solutions. They prioritized energy efficient resources over their alternatives where possible. They also seemed to have paid attention to the configuration of their processors to achieve energy efficiency. With most of the processors configurable, students had the option to make changes to the configurations to suit the intended separation operation, which leads to more energy-efficient solutions. For instance, to separate a salt solution, students could configure a boiler to 100^{o}C which is energy efficient for the required task (boiling temperature of water is 100^{o}C) or leave it at the default setting of 200^{o}C which will still separate the solution but with higher energy consumption. Interestingly, 72% of the solutions of the students fall into this category which is a good indication that majority of the students viewed the game tasks from an engineering point of view and thought through their solutions as expected.

The other two clusters of students did not perform as well as the Optimal Solution group. However, the solutions patterns of these students only account for 28% of the total. Nonetheless, the solution of the Optimal Energy group was very similar to those of the Optimal Solution group discussed above. As can be seen in Table 5, students in this group expended far less energy compared to the Optimal Resource group, but a little more than the Optimal Solution group. When compared to the other two groups, the Optimal Energy students used the most amount of resources when solving the given tasks. This indicates that although these students may have configured and sequenced their processors accurately to achieve energy efficiency, they used far higher number of resources than was necessary. The Optimal Resource group, on the other hand, used the least amount of resources compared to the other two groups. However, their solutions consumed a significantly higher amount of energy. This implies that students in this group used fewer resources to achieve the necessary separations, but they probably used the least energy efficient processors and also failed to configure their processors to achieve energy efficiency. Overall, the solution pattern of the Optimal Resource group could be considered the worst of all three solution patterns obtained.

3.3 Solution Pattern Characteristics

To better understand the three solution patterns obtained from the gameplay data, the mean scores of the considered variables were plotted against the cluster groups. First, a look at the trend of energy consumption across all 25 levels shows the similarity between the Optimal Energy group and the Optimal Solution group, as seen in Figure 5. It also shows that overall, the Optimal Resource group did not deviate systematically, but consumed significantly more energy in some game levels, with a significant spike at level 20. All four solutions in this group used an average of over 1200 units of energy on this level compared to an average of about 90 units used by the other groups. The game task for this level required students to separate 65 units of iron and a mould containing a mixture of one unit of iron and four units of bricks. The optimal sequence of solution used by most of the students employed first a sieve to separate the 65 units of iron, next a shredder to shred the mould of iron and bricks into single units, and finally, a magnet separator to separate the iron from the bricks. On the other hand, the sequence solution used by all the students in the Optimal Energy group involved a shredder first to shred all the materials before passing these through a magnet separator. This is an energy intensive process given that the energy consumed by a shredder and a magnet separator to process one unit of material is higher than that required by a sieve. With all the materials for separation passed through the shredder and the magnet, much more energy is consumed than when a sieve is first used to separate a bulk of the materials. This highlights the limitation of the sequential solution pattern of the Optimal Resource group. It also provides insight into some misconceptions of students about the use of the shredder and best practices when separating materials of different volumes. This would potentially help educators to design specific game tasks to assess specific learning outcomes and provide tailored feedback to the affected students.

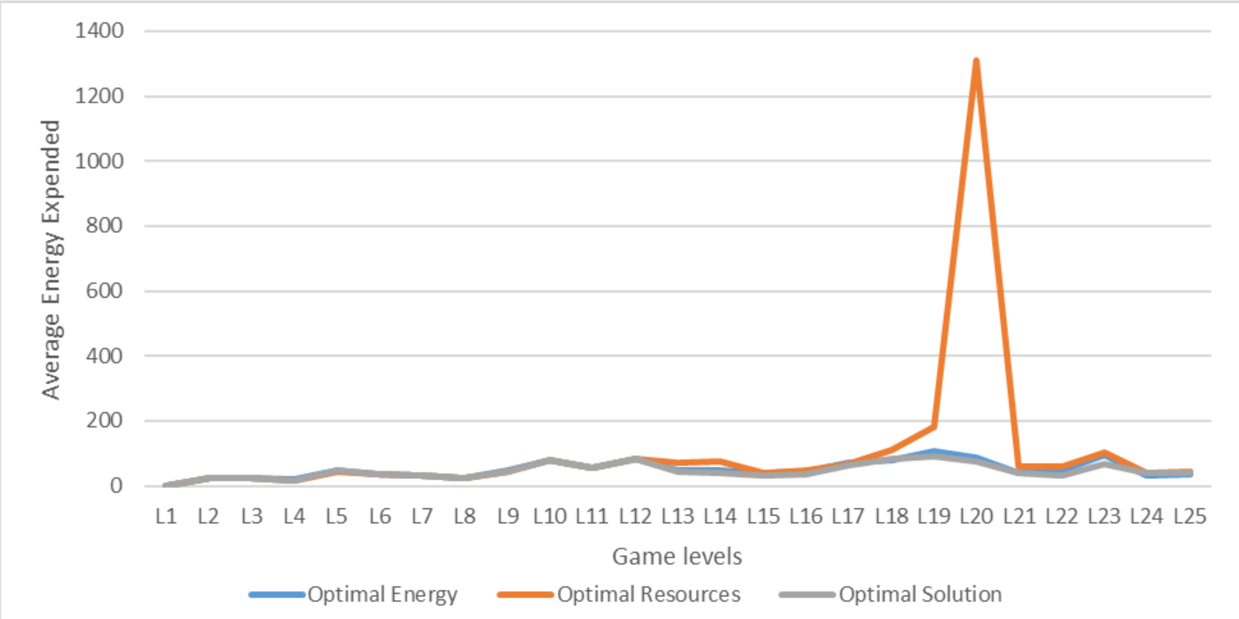

Figure 5: Visualization of energy consumption across game levels by cluster membership.

Lastly, Figure 6 provides better insight into the number of resources used by the different groups across the 25 levels. As initially seen in Table 5, the trend shows that the Optimal Energy group used more resources on average across the levels compared to the other two groups.

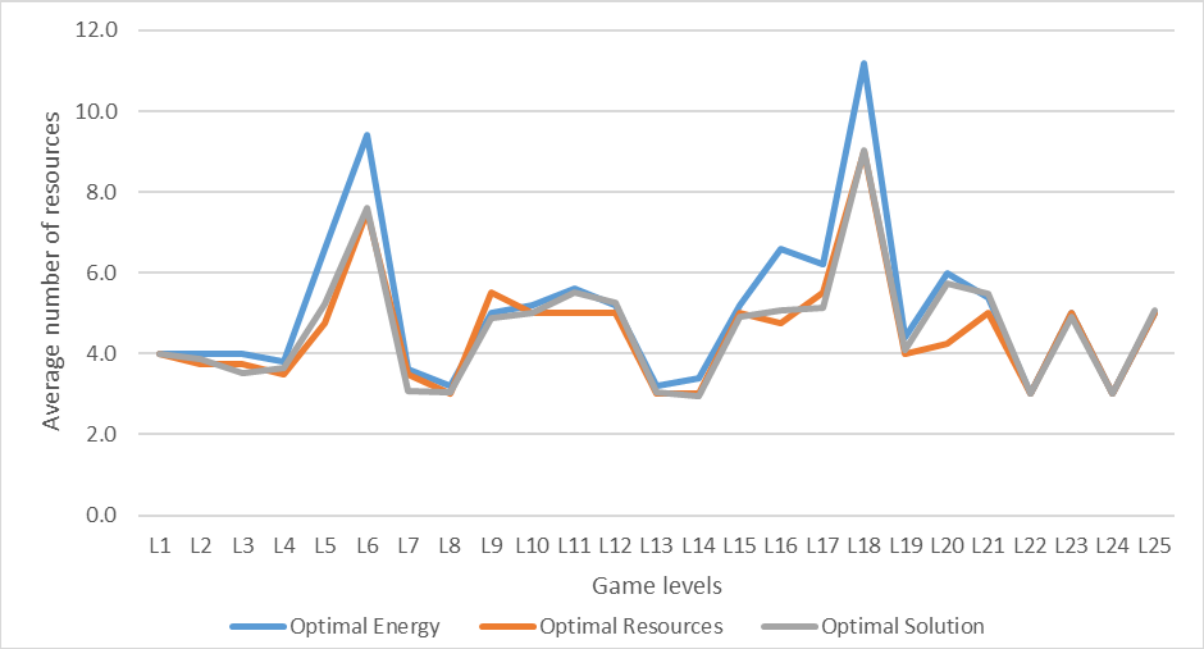

Figure 6: Visualization of resources used across game levels by cluster membership.

Spikes in the number of resources used by all three groups were seen in Levels 6 and 18. Generally, the solution sequences used by the students showed that they used an excessive number of conveyors and sieves compared to what was required for solving those levels. Although the energy expended from the use of conveyors and sieves is negligible in this game, in reality, this may not always be the case, as having unnecessary process units in place does have significant implication in terms of space, capital costs and sustainability footprints. The visual representations of solution patterns obtained from game log data provide educators as well as researchers better views into gameplay behaviour patterns of students. These results provide insights into the strengths and weaknesses of each student and show how log data of gameplay can be used to better understand game performance in order to develop in-game assessment strategies and provide relevant and tailored feedback.

4 CONCLUSIONS

This study builds upon our previous research examining the relationship between gameplay performance, game experiences and the perceptions of students towards educational games. In this study, the behaviour, performance and solution patterns of students were examined. Using game log data, it was found that there was high variability in the sequential solutions used by the students across the first levels of the game. This highlighted the exploratory behaviour of students when presented with an unfamiliar game environment. This variability in solutions was seen to diminish over time as students got used to the game mechanics. This finding is considered interesting as it offers insight to game developers, educators, and researchers into what to expect when deploying games for learning purposes. In particular, it highlights the relevance of pre-game training to ensure that all students get used to the game environment before an intervention study. It also points to the need for extended gameplay time when using games for assessment purposes.

The gameplay solutions patterns of students showed the different strategies used to solve the game tasks. Over 70% of the students were categorised as having the most sustainable solutions (Optimal Solution group). A closer look at the sequential solution patterns provided insight into the type, sequencing, and number of resources used by the students. While the performance of 13% of the students was considered poor, the solution patterns of the rest of the students were commendable. The results indicate that students thought through their solutions and applied their knowledge of sustainability during the process.

These results show how much information can be obtained from game log data and how useful they could be for conducting performance-based assessments. Although tools to automate processing and analysing log data need to be further developed, the information obtained from these data can rarely be obtained from other forms of assessment. Future studies could look at developing simpler and structured ways of logging and processing game log data. While this study provides evidence of the behaviours and performance of students in a learning game, additional studies with larger sample sizes are required to verify these findings.

Acknowledgements: Thanks are due to Prof Tom van Gerven of KU Leuven, Belgium and Dr Vijesh Bhute of Imperial College, London for their support throughout the data collection period.

Declaration: This project has received funding from the European Union's EU Framework Programme for Research and Innovation Horizon 2020 under Grant Agreement No 812716

REFERENCES

Almond, R. G. (2015). Tips and Tricks for Building Bayesian Networks for Scoring Game-Based Assessments. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 9505, pp. 250–263). https://doi.org/10.1007/978-3-319-28379-1_18

Bellotti, F., Kapralos, B., Lee, K., Moreno-Ger, P., & Berta, R. (2013). Assessment in and of Serious Games: An Overview. *Advances in Human-Computer Interaction*, *2013*, 1–11. https://doi.org/10.1155/2013/136864

Cohen, L., Manion, L., & Morrison, K. (2017). Factor analysis, cluster analysis and structural equation modelling. In *Research methods in education* (8th Editio, pp. 818–838).

Creswell, J. W. (2011). *Educational Research* (4th ed.). Retrieved from http://basu.nahad.ir/uploads/creswell.pdf

Glassey, J., & Haile, S. (2012). Sustainability in chemical engineering curriculum. *International Journal of Sustainability in Higher Education*, *13*(4), 354–364. https://doi.org/10.1108/14676371211262308

Glassey, J., & Magalhães, F. D. (2020). Virtual labs – love them or hate them, they are likely to be used more in the future. *Education for Chemical Engineers*, *33*, 76–77. https://doi.org/10.1016/j.ece.2020.07.005

Groff, J. S. (2018). The potentials of game-based environments for integrated, immersive learning data. *European Journal of Education*, *53*(2), 188–201. https://doi.org/10.1111/ejed.12270

Hou, H.-T. (2012). Exploring the behavioral patterns of learners in an educational massively multiple online role-playing game (MMORPG). *Computers & Education*, *58*(4), 1225–1233. https://doi.org/10.1016/j.compedu.2011.11.015

Kang, J., Liu, M., & Qu, W. (2017). Using gameplay data to examine learning behavior patterns in a serious game. *Computers in Human Behavior*, *72*, 757–770. https://doi.org/10.1016/j.chb.2016.09.062

Kerr, D., & Chung, G. K. W. K. (2012). Identifying Key Features of Student Performance in Educational Video Games and Simulations through Cluster Analysis. *JEDM - Journal of Educational Data Mining*, *4*(1), 144–182. Retrieved from http://www.educationaldatamining.org/JEDM/index.php/JEDM/article/view/25

Kim, Y. J., Almond, R. G., & Shute, V. J. (2016). Applying Evidence-Centered Design for the Development of Game-Based Assessments in Physics Playground. *International Journal of Testing*, *16*(2), 142–163. https://doi.org/10.1080/15305058.2015.1108322

Loh, C. S., & Sheng, Y. (2015). Measuring the (dis-)similarity between expert and novice behaviors as serious games analytics. *Education and Information Technologies*, *20*(1), 5–19. https://doi.org/10.1007/s10639-013-9263-y

Loh, C. S., Sheng, Y., & Ifenthaler, D. (2015). Serious Games Analytics: Theoretical Framework. In *Serious Games Analytics* (pp. 3–29). https://doi.org/10.1007/978-3-319-05834-4_1

Loh, C. S., Sheng, Y., & Li, I.-H. (2015). Predicting expert–novice performance as serious games analytics with objective-oriented and navigational action sequences. *Computers in Human Behavior*, *49*, 147–155. https://doi.org/10.1016/j.chb.2015.02.053

Moseley, A. (2013). A Case for Integration. In T. Connolly, T. Hainey, E. Boyle, G. Baxter, & P. Moreno-Ger (Eds.), *Psychology, Pedagogy, and Assessment in Serious Games* (pp.

342–356). https://doi.org/10.4018/978-1-4666-4773-2.ch016

Perini, S., Oliveira, M., Margoudi, M., & Taisch, M. (2018). The Use of Digital Game Based Learning in Manufacturing Education – A Case Study. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*: *Vol. 10925 LNCS* (pp. 185–199). https://doi.org/10.1007/978-3-319-91152-6_15

Shute, V. J., & Ke, F. (2012). Games, learning, and assessment. In *Assessment in Game-Based Learning: Foundations, Innovations, and Perspectives* (pp. 43–58). https://doi.org/10.1007/978-1-4614-3546-4_4

Shute, V. J., Wang, L., Greiff, S., Zhao, W., & Moore, G. (2016). Measuring problem solving skills via stealth assessment in an engaging video game. *Computers in Human Behavior*, *63*, 106–117. https://doi.org/10.1016/j.chb.2016.05.047

Smith, S. P., Hickmott, D., Southgate, E., Bille, R., & Stephens, L. (2016). Exploring Play-Learners' Analytics in a Serious Game for Literacy Improvement. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*: *Vol. 9894 LNCS* (pp. 13–24). https://doi.org/10.1007/978-3-319-45841-0_2

Suescún, E., Camapio do Prado Leite, J. C., Werneck, V., Mazo, R., Vallejo, P., Toro, M., … Rincón, R. (2018). SimulES-W: A Collaborative Game to Improve Software Engineering Teaching. *Computación y Sistemas*, *22*(3). https://doi.org/10.13053/cys-22-3-2711

Thanasi-Boçe, M. (2020). Enhancing students' entrepreneurial capacity through marketing simulation games. *Education + Training*, *ahead-of-p*(ahead-of-print). https://doi.org/10.1108/ET-06-2019-0109

Udeozor, C., Russo Abegão, F., Gill, S., & Glassey, J. (n.d.). *Relationship between Perceptions and Performance of Students in a Serious Game.*

Udeozor, C., Toyoda, R., Russo Abegão, F., Gill, S., & Glassey, J. (submitted). Perceptions of the Use of Virtual Reality Games for Chemical Engineering Education and Professional Training. *Higher Education Pedagogies*.

Westera, W., Nadolski, R., & Hummel, H. (2014). Serious Gaming Analytics: What Students´ Log Files Tell Us about Gaming and Learning. *International Journal of Serious Games*, *1*(2). https://doi.org/10.17083/ijsg.v1i2.9

Wouters, P., & van Oostendorp, H. (2013). A meta-analytic review of the role of instructional support in game-based learning. *Computers & Education*, *60*(1), 412–425. https://doi.org/10.1016/j.compedu.2012.07.018

Xenos, M., & Velli, V. (2020). A Serious Game for Introducing Software Engineering Ethics to University Students. *International Conference on Interactive Collaborative Learning*, 579–588. https://doi.org/10.1007/978-3-030-11932-4_55

Yue, W. S., & Tze, H. Y. (2015). Effectiveness of MMORPGS in Enhancing Student Interaction. *2015 IEEE 15th International Conference on Advanced Learning Technologies*, 152–154. https://doi.org/10.1109/ICALT.2015.60