Is Gamification an effective pedagogic strategy? Using *Cuppa Services* to teach lean and process improvement

Nicholas Wake
Nicholas.Wake@faculty.hult.edu
Hult International Business School

Alexander Urquhart
Hult International Business School

Corina Kroon

Abstract

There is an increasing body of literature that discusses the benefits of game playing – for example Vlachopoulos and Makri (2017) present a systematic literature review of the cognitive, behavioural and effective outcomes of games and simulations.

In this paper an updated version of the Jackson (1996) *Cups Game* is presented as an approach to teaching lean and process improvement. The research correlates the student-self-reported experience of the game with their ability to recall key lean and process improvement concepts. The mediating effect of experience on student key concept recollection is discussed.

The paper reflects on the advantages of using this game to teach lean and process improvement.

Keywords: teaching, pedagogy, simulation

Introduction

From the very beginning, the lecture has been at the heart of teaching in universities. Increasingly a number of other pedagogic strategies are being used to teach operations and supply chain management. For example, Lewis and Maylor (2007) discuss experiential learning approaches in which knowledge co-production occurs and Brandon Jones et al. (2012) discuss a number of approaches to teaching, including simulations, virtual learning environments, cases and role-play.

Whilst there is an increased use of games (the term 'gamification' is often used to capture these concepts), using games pedagogically can only be justified where evidence-based research shows they are effective. Prior research (Wake and Walley, 2016) suggests that teaching faculty overwhelmingly use social processes (peer observation, trial and error and reflection) and not the pedagogic literature to develop teaching practice. Thus whilst memetic isomorphism may account for increased gamification of teaching, evidence that gamification is an effective strategy for teaching operations management is limited.

This paper reports the results of research in which students were taught the principals of lean and process improvement through the business game *Cuppa Services*. The paper discusses the correlation between students' ability to recall key learning points and student self-reported measure of game effectiveness, subject matter understanding, and the likelihood that they recommend the game to others.

Meaningful Games

In its purest form, game play has been defined as:

> *a free and meaningful activity, carried out for its own sake, spatially and temporally segregated from the requirements of practical life, and bound by a self-contained system of rules that holds absolutely* (Huizinga 1949).

This definition is a useful point of entry into the discussion of play in that, for a game to have pedagogic benefits, it must be 'meaningful', linking to the curriculum and thus directed towards defined learning outcomes. Games are fun and stimulating (Brown and Vaughan, 2009); as previous theoretical (Wake and Johnston, 2011) and empirical (Urquhart and Wake, 2017) research shows, positive emotions lead to better learning outcomes.

The literature discusses the pedagogic benefits of game play. For example, Kapp (2012) argues that games allow participants to acquire and retain skills and knowledge with high retention rates, accelerating the learning curve.

This research tests the hypothesis that using the *Cuppa Services* simulation will create understanding of lean and process improvement principles, as well as an effective learning environment that leads students to better recall the key learning points of these topics (Figure 1).

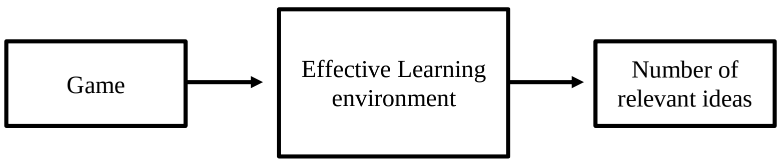

Figure 1

In understanding the effectiveness of the game, the mediating effect of experience will be examined to see if it is a simulation that is better suited to participants with significant, some, or no experience with lean and process improvement. Additionally, the research aims to see whether it is effective at teaching the lean and process improvement tools, or whether it is more suited to teaching the 'softer' philosophy elements of implementation.

Methodology

For this research, two executive MBA cohorts participated in a simulation called Cuppa Services, which is a heavily modified version of the original *Cups Game* (Jackson, 1996). In the Jackson (1996) version, 'interested spectators' observe a five-stage assembly process. Through three rounds of simulation the process is changed, allowing the instructor to 'demonstrate' the difference between push- and pull-based systems, as well as the effect on inventory of halving batch size.

A similar five-stage process is used in *Cuppa Services*, however all students are allocated roles—logistics, worker, quality control, manager, customer or process improvement consultant—and so are actively involved in the simulation. Space precludes a more detailed description of the game, but a full explanation is available at www.operations.university.

The key difference between this game and the Jackson (1996) version is that inductive learning takes place: Through five rounds of simulation, participants record productivity, quality, work in progress, throughput time and on-time delivery data. These data are discussed and the students are challenged to come up with and implement their own process improvement ideas, which are incorporated into the next round of simulation. To draw on Kolb (1984), the simulation gives students a concrete experience of running a process. They are asked to reflect on the experience and, through a process of abstract conceptualising, identify improvements that they actively experiment with. These ideas are discussed and validated with the instructor, who will label their ideas using the language of lean (e.g., "Your idea to do X is called a poka-yoke" or "You are suggesting something which is referred to as 'single piece flow'").

Details are included to increase the fun environment of game play: Participants are asked to wear hair nets because they are working in a food factory that requires high standards of hygiene; the artefacts of play (game boards, etc.) are brightly coloured cartoons (see Appendix A); and the reviews that occur between rounds are injected with humour laughter and enjoyment.

At the end of play (which takes around three hours) participants are required to break into small groups to reflect on what they learned, presenting their ideas in plenary. The instructor brings the discussion to a close, drawing on lean and process improvement principals to provide a narrative and summary of the learning. For the purposes of this research a checklist of lean and process improvement principles is used to ensure that the cohorts receive the same information.

After the simulation, participants were invited to fill in a questionnaire that rated their satisfaction with the game (six variables on a five-point Likert scale), their understanding of the subject, and their experience of lean and process improvement. The survey asked if they could recall and if they found relevant a list of 14 lean and process improvement ideas that the game was designed to teach. A total of 97 usable questionnaires were returned from a combined class size of 137 students.

The *Cuppa Services* simulation has been a standard and anecdotally effective part of the syllabus for this Operations Management course for several years. Due to the importance of the student experience and the priority of quality student learning, however, approval to distribute the survey was obtained from both the campus dean and the school Research Ethics Committee. Furthermore, participants were advised that participation in the survey was completely optional.

Analysis

Experience

Prior experience with lean and process improvement principles varied throughout the group: 45.1% of participants indicated that they had no experience with lean or process improvement; 42.8% of the group had some experience; only 12.1% of the group indicated significant experience.

Satisfaction

In general, the simulation was considered to be effective: 65.6% of participants strongly agreed that the simulation was an effective approach to teach the subject, and 60.2% strongly agreed that they would recommend the approach to other students.

Remember and Relevance

In the post-simulation questionnaire, participants were asked which concepts they recognized from the activity, and the extent to which they found these concepts relevant to their own business. Table 1 presents full results for all 14 concepts.

Out of the 14 concepts taught, on average, participants remembered ten (*SD* = 3.44), with a low of two ideas and a high of all 14. The number of concepts that participants found relevant to their business was lower, with an average of eight (*SD* = 3.15), a low of 0, and a high of all 14.

The most commonly remembered concept was the ‘Role of leadership’, where 90.0% of the sample remembered the idea. This was followed by ‘Involve everyone’ with 88% recall. ‘Plan, do, check, act’ was the concept remembered by the fewest participants (68%).

Participants found the role of leadership as the most relevant concept (75%), followed by continuous improvement (71%). The least relevant were clearly ‘Single piece flow’ (25%) and ‘kanbans’ (34%).

Table 1 – Post-simulation questionnaire results

Learning concept present in simulation	**Remembered after simulation (%)**	**Relevant to own business (%)**
Role of leadership	90	75
Involve everyone	88	69
Waste elimination	86	64
Sand cone model	85	58
Voice of customer	85	69
Continuous improvement	85	71
Inventory reduction	85	50
Throughput time reduction	79	44
Supply chain synchronisation	75	42
Kanbans	74	34
Push vs. pull	74	40

Define, measure, analyse, improve, control	74	53
Single piece flow	71	25
Plan, do, check, act	68	49

Satisfaction inferential
The six Likert satisfaction variables were highly negatively skewed (the majority of the participants rated the same) which is problematic for linear analysis. Thus the variables were recoded and a log transformation performed to change the scores to be normally distributed. In the recoded variables, a low score is positive and a high score is negative. These transformed scores were included in a Pearson's correlation analysis which finds all satisfaction variables significantly correlated with each other.

The positive significant correlations indicate that the variables are related to each other. For example, participants who found the simulation 'Enjoyable' also rated it as highly 'Effective'. Due to this, a scale was created out of this named 'Satisfaction'. A reliability analysis of these six items indicated a good internal consistency (α = 0.80), as shown in Table 2.

Table 2 – Pearson's correlation analysis of satisfaction results
** = p < .05, ** = p < .01, *** = p < .001*

	1. Enjoyable	2. Effective	3. Understandable	4. Applicable	5. Recommend	6. Length of time
1.	-					
2.	.57***	-				
3.	.43***	.49***	-			
4.	.30**	.23*	.43***	-		
5.	.53***	.47***	.42***	.38***	-	
6.	.37**	.23*	.35**	.33**	.54***	-

The role of experience
The role of experience had little impact on how participants rated the simulation. The 11 participants with significant experience rated the simulation slightly better than the other groups (*M* = 1.59, *SD* = 0.52), while the participants with no experience (*M* = 1.64, *SD* = 0.42) or some experience (*M* = 1.65, *SD* = 0.50) rated the simulation nearly the same. These minor differences were not statistically significantly different, according to a one-way ANOVA test (*f*(2, 88) = 0.11, *p* = .90).

A comparison was made between the two experience groups concerning their ability to recall the key learning points and their perception of the relevance of these points to their business. The participants with some experience (*M* = 10.33, *SD* = 3.78) remembered a similar amount of ideas from the simulation compared to the participants with no experience (*M* = 10.32, *SD* = 3.10). People with significant experience remembered fewer ideas compared to the other groups (*M* = 8.63, *SD* = 3.67). However, these differences were not significantly different according to a one-way ANOVA test (*f* (2, 88) 1.14, *p* = 0.32). Furthermore, participants with some experience found the most ideas relevant to their business (*M* = 8.90, *SD* = 3.24), followed by participants with no experience (*M* = 7.51, *SD* = 2.85); participants with significant experience (*M* = 7.10, *SD*

= 2.66) found the fewest ideas to be relevant. Again, however, these differences were not statistically significant (f (2, 88) = 2.77, p = .07).

Effectiveness of game in supporting learning outcomes

The 14 key learning points in Table 1were allocated to one of three categories: ‘Philosophy’ (sand cone model, inventory reduction, waste elimination, continuous improvement, supply chain synchronisation), ‘Tools/Techniques’ (kanbans, push vs. pull, throughput time reduction, single piece flow, DMAIC and PDCA), and ‘Soft Skills’ required for process improvement (involving everyone, role of leadership, listening to the voice of the customer). This categorisation follows Slack and Brandon-Jones (2018), who discuss the philosophy of lean (p. 409), some of the tools and techniques (p. 393) and the requirement for leadership (p. 447).

Participants remembered different proportions of the idea groups (‘Philosophy’, ‘Tools’, and ‘Soft’). These findings were significant: Participants remembered on average 94% of ‘Soft’ concepts, 89% of ‘Philosophy’ concepts, and 79% of ‘Tool’ concepts.

The questionnaire scores for ‘Remembered’ concepts were analysed with a repeated measures ANOVA. Mauchly’s test indicates that the assumption of sphericity is violated (x^2 (2) = 14.98, p < .01). Using a Greenhouse-Geisser corrected test (ε = .87), the findings suggest that participants remembered different proportions of the concept categories (f(1.74, 159.75) = 28.01, p < .001, ηp^2 = .238). A follow-up Bonferroni pairwise comparisons analysis of the groups indicates that the proportion of ‘Soft’ concepts remembered is statistically significantly higher than the ‘Philosophy’ concepts (d = 0.33, p = .005) and the ‘Tool’ concepts (d = 0.73, p = .000). Additionally, the proportion of remembered ‘Philosophy’ items is statistically significantly higher than the ‘Tool’ concepts (d = 0.45, p = .000).

The questionnaire scores for ‘Relevant’ concepts were similarly analysed. The assumption of sphericity is violated (x^2 (2) = 13.91, p < .01); a Greenhouse-Geisser corrected test (ε = .88) finds that participants remembered different proportions of the idea groups (f(1.75, 161.16) = 43.99, p < .001, ηp^2 = .323). Follow-up Bonferroni pairwise comparisons analysis of the groups indicates that the proportion of the ‘Soft’ concepts remembered is statistically significantly higher than the ‘Philosophy’ concepts (d = 0.49, p = .000) and the ‘Tool’ concepts (d = 1.05, p = .000). Additionally, the proportion of remembered ‘Philosophy’ concepts is statistically significantly higher than the ‘Tool’ concepts (d = 0.56, p = .000.)

Experience and learning outcomes

There are some indications that experience has some effect in what the participants remembered. Table 3 shows the findings of three tests, including mean differences between each concept group. A mixed factorial ANOVA (where the dependent variables are ability to ‘Remember’ the three concept categories, and the independent variable is ‘Experience’) was used to analyse whether there are differences between experience groups.

Table 3 – Results from three repeated measures ANOVA tests
* = $p < .05$, ** = $p < .01$, *** = $p < .001$

	df	*f*	*p*	ηp^2	Philosophy – Soft	Philosophy – Tool	Soft – Tool
No experience	1.63, 65.01	15.4	.000	.278	-9.11**	10.41*	19.51***
Some experience	1.61, 61.25	7.72	.002	.169	0.51	8.63**	8.12*
Lots of experience	2, 20	12.99	.000	.565	-16.36*	9.39	25.76**

The mixed factorial ANOVA indicated a significant interaction of 'Remembering' and 'Experience' ($f(3.3, 145.08) = 3.30$, $p < .05$, $\eta p^2 = .007$). A one-way ANOVA per concept category between 'Experience' was conducted and three repeated measures ANOVA, one per experience group were calculated per group of ideas.

The between group one way ANOVA indicated that there was a significant effect of 'Experience' on remembering 'Philosophy' concepts ($f(2, 90) = 6.98$, $p < .01$). That is, the participants with some experience remembered a significantly larger proportion of the 'Philosophy' concepts than the group with lots of experience ($d = 0.97$, $p < .01$). However, there was no significant difference between the group with no experience and the group with some experience. Additionally, there was no significant difference between the group with lots of experience and the group with no experience.

No significant difference between the experience groups is identified for the 'Soft' concept category ($f(2, 90) = .64$, $p < .05$). But there is a significant difference between the experience groups based on the proportion of remembered 'Tool' concepts ($f(2, 90) = 3.67$, $p < .05$). Similar to 'Philosophy', the group with some experience remembered significantly more 'Tool' concepts than the group with lots of experience ($d = 0.79$, $p < .05$). There was no significant difference between the group with no experience and the group with some experience. Additionally, there was no significant difference between the group with no experience and the group with lots of experience.

The role of satisfaction

A correlation analysis was made between both the combined 'Satisfaction' variable and all individual variables with the count of 'Remembered' ideas and the count of 'Relevant' ideas. The overall satisfaction was not related to how many ideas they remembered or found relevant. However, finding the simulation effective and understanding the simulation was significantly related to finding more ideas relevant ($r(\text{n} = 93) = -.21$, $p < .05$ and $r(\text{n} = 92) = -.23$, $p < .05$). Furthermore, participants that were more likely to recommend the simulation found more ideas relevant (r (n = 93) = $-.23$, $p < .05$). This finding is presented diagrammatically in Figure 2, and is different from the correlation hypothesised in Figure 1.

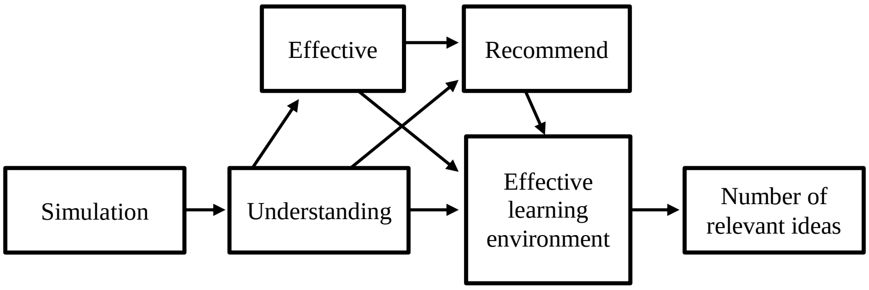

Figure 2 – Diagram of gamification, representing interrelationships of concepts impacting students' experience of the Cuppa Services *game*

The overall satisfaction scale was not related to findings the ideas relevant but understanding the simulation and finding it effective was.

Discussion

Concepts remembered

The data show that participants remembered on average 94% of 'Soft' concepts, 89% of 'Philosophy' concepts, and 79% of 'Tool' concepts. This is an interesting finding, and perhaps is a reflection of a teaching strategy that involves experiential learning and deduction. Perhaps this pedagogic strategy is a more powerful way to demonstrate the role of leadership (and other 'Soft' concepts) than a lecture—an argument that supports the notion that the simulation is effective.

Relevance

Compared with concepts remembered, the relevance of the of each of the 14 aspects taught was lower. Single piece flow (which 25% of the class found relevant) and kanbans (34%) were the lowest scoring items on the list. It is worth considering why this is. It could be that the participants were able to accurately assess that these techniques lacked relevance to their industry, or it could be that the simulation was not effective in explaining the technique in a way that allowed abstract conceptualisation and application to a new context. Alternatively, it could be that determining the 'relevance of a concept is, to use the Bloom (1956) taxonomy cognitively more difficult to do since it requires evaluation of concepts and not simply remembering them. There is insufficient evidence to postulate the cause of this, however the research also captured qualitative data which will be analysed for a future paper, and this might provide a narrative on why this is. It should be noted that within the context of a three to four-hour simulation, the level appears to be high.

Effect of Experience

It is interesting to note that all participants, irrespective of experience, rated the simulation as effective. This suggests that the game is appropriate for a broad cross section of participants. When considered individually, it is interesting that each experience group remembered the same number of learning points, but that there are differences in the number of learning outcomes remembered when they are clustered. In summary, the analysis shows that:

1. Participants with some experience remembered a larger proportion of philosophy items compared to the group with lots of experience. However, there was no difference between some and no experience or lots and no experience on the proportion of philosophy items.
2. The experience groups did not differ on soft items.
3. The group with some experience remembered a significantly larger proportion of tool items compared to the group with lots of experience. But there was no difference between some and no experience or lots and no experience on the proportion of tool items.

Intuitively, one would expect those with most experience to remember most of the items, i.e. familiarity with lean and process improvement techniques would be an advantage. However the data shows that for the philosophy and tool items, the group with some experience remembered more items. Could it be that experience became a barrier to learning? Information discussions with the participants revealed that those with significant experience had a range of experience and included people with green, black, and master black belt statuses.

Conclusions

The data suggests that the participants were very satisfied with the game, and that it was an effective pedagogic strategy for allowing students to remember the concepts highlighted. The literature (e.g. Wake and Johnston 2011) demonstrates that teaching is most effective when students experience positive emotions, and a game has the benefit of both removing anxiety and boredom from the classroom whilst simultaneously making learning ‘fun’. It appears that this game is successful in doing this as are other strategies such as ‘hooks’ (Brandon-Jones, 2011), and other innovations (Brandon-Jones et. al., 2012).

This paper advances our understanding of the effectiveness of game play since it goes beyond simply relying on student self-reported data on satisfaction but instead links this to their ability to remember the key concepts, and also looks at the impact of experience.

Work (e.g. Reimann and Junge, 2017) is being undertaken to evaluate supply chain games, but there is little data on which games are used regularly in teaching operations management and supply chain management. A study to review which games are used would reveal the extent to which gamification is being adopted as a pedagogic strategy amoungst OM faculty. . The work of Reimann and Junge (2017) is a start since it provides a comprehensive list of the games that appear in the academic literature (but thus does not catalogue any ‘home grown’ games). This work is to be encouraged if we, as an association of operations management professionals and academics, are to inspire the next generation through our teaching practice.

References

Bloom, B.S. (1956), The Taxonomy of Educational Objectives, the Classification of Educational Goals, Susan Fauer Company, New York, N.Y.

Brandon-Jones, A., 2011, Round-table presentation: Hooks in Teaching, European Operations Management Association

Brandon-Jones, A., Piercy, N., Slack, N., 2012, Bringing teaching to life: Exploring innovative approaches to operations management education, International Journal of Operations & Production Management, Vol. 32 Issue: 12, Brown, S., Vaughan, C., 2009, *Play: How it Shapes the Brain, Opens the Imagination, and Invigorates the Soul*, Penguin, New York

Huizinga, J., 1949, *Homo Ludens*, London, Routledge

Jackson, P., 1996 , *The Cups Game*, NSF Product Realisation Consortium Nodule Description, Cornell University

Kapp, K. M. 2012. The gamification of learning and instruction: game-based methods and strategies for training and education, John Wiley & Sons.

Kolb, D.A., 1984, *Experiential Learning: Experience as the Source of Learning and Development*. Prentice-Hall, Englewood Cliffs

Lewis, M., Maylor, H., 2007, *Game playing and operations management education*, International Journal of Production Economics, 105: 1

Reimann, F., Junge, A., 2017, *Deceloping a systematization framework and evaluation of supply chain games*, Conference Proceedings, European Operations Management Association

Slack N., Brandon-Jones, A., 2018, *Operations and Process Management: Principals and Practice for Strategic Impact*, Pearson, Harlow

Urquhart, A., Wake, N., 2017, *The Effects Of Teaching Style On Student Recall*, Conference Proceedings, European Operations Management Association

Vlachopoulos, D., Marki, A., 2017, *The effect of games and simulations on higher education: a systematic literature review*, International Journal of Education Technology in Higher Education, 14: 22

Wake, N., Johnston, B., 2011, *Designing Effective Approaches to Teaching Operations Management*, Conference Proceedings, European Operations Management Association

Wake N., Walley, P., 2016, *Learning to Teach Operations*, Conference Proceedings, European Operations Management Association

Appendix A: Example of artefacts used to play Cuppa Services