

Examining the Effect of Gamification in Information Science, Computer and Engineering Education: A Meta-analysis of Student Learning Performance

Shurui BAI ^{a*}, Khe Foon HEW ^b & Biyun HUANG ^c
Faculty of Education, The University of Hong Kong, Hong Kong SAR
*shurui18@hku.hk

Abstract: Gamification is commonly defined as the use of game design elements in non-game contexts. Compared with the traditional lecture-based class, a gamified class is commonly seen as more engaging and joyful. Anecdotal reports suggest that students perceive the use of gamification positively. But does gamification really improve student learning performance? The conflicting results reported in previous studies make the decision to support (or to dismiss) the use of gamification in education difficult. This meta-analysis examined the overall effect size of gamification on student learning performance in the contexts of information science, computer and engineering education. Thirteen studies employing between-subject designs that compared the effects of gamified versus non-gamified courses on student learning performance constituted the current sample. Results using the random-effect model revealed a significant small effect of gamification on students' learning achievement (Hedges's $g = 0.36$, $CI = 0.006 - 0.714$, $p = 0.046$). Results of various moderator analyses were discussed. Suggestions regarding the selection and use of the game design elements were provided. We conclude that the use of gamification results in higher learning gains compared with the non-gamified courses within the contexts of information science, computer and engineering education.

Keywords: Gamification, information science, computer and engineering, learning performance, meta-analysis

1. Introduction

Gamification is widely described as the use of game design elements in non-game contexts (Deterding, Dixon, Khaled, & Nacke, 2011). Game design elements are the basic building components of gamification and the common game design elements include badges, challenges, leaderboard/rank, levels/unlock, storyline, points, progress bar (Sailer et al., 2017).

Gamification shows a natural relation with information science, which has developed from an entertainment-oriented technology only for productive aims into mainstream in information science domain (Koivisto & Hamari, 2019). Despite the increasing demand of the market, software engineering and computer science disciplines (e.g., programming) present a high dropout rate in schools (Narasareddygari, Walia, & Radermacher, 2018).

Hitherto, recent reviews show that empirical studies on gamification in information and computer science education have markedly surpassed other subject disciplines (Dicheva, Dichev, Agre, & Angelova, 2015; Subhash & Cudney, 2018). However, findings that emerged from the use of gamification in the field of information, computer and engineering education have been mixed. For example, on one hand, Marín et al. (2019) stated that the effect of gamification to teach C programming is encouraging since students earned higher marks when the gamified platform was used. On the other hand, De-Margos et al. (2017) reported that gamification failed to improve student learning compared to the control group. Students in the control (non-gamified) group obtained significantly higher scores in the final exams compared to the gamified group. This therefore leads us to an important question: "*Does gamification indeed improve students learning outcomes or it is merely a hype*"?

2. Literature Review

2.1 Previous Reviews

In this section, we briefly reviewed information, computer and engineering education literature in recent years (2010-2018). So far, only several narrative syntheses have been published.

Souza, Veado, Moreira, Figueiredo, and Costa (2018) conducted a systematic mapping study to identify what game-related methods had been used to support software engineering education. Leaderboards, points, and levels were found to be the most frequently used game elements. No quantitative analyses of effect sizes were reported.

Gari, Walia, and Radermacher (2018) examined 16 studies in computer programming and software engineering education. The researchers found that badges, points and leaderboard were the most commonly used game elements. The main findings of each study were summarized. However, similar to other previous reviews, this review did not present any quantitative analyses of the studies' effect sizes.

Narasareddygari et al. (2018) reviewed 16 empirical studies in software engineering and computer science education disciplines. The authors counted the frequency of most used game elements and summed up the benefits of gamified learning in each study, then some suggestions for the design of learning environment were provided. However, neither effect sizes to substantiate the effectiveness of gamification nor synthesis of key elements for a successful gamified class were presented. The review just summarized results from each study.

2.2 Contribution of the Present Study

We contributed the present meta-analysis examining the effectiveness of gamified class on student academic performance compared to non-gamified class. A meta-analysis, which integrates the results of several independent small studies, can provide a more precise estimate of the effect of an intervention. Meta-analyses can also determine the strength of the effect, establish whether the direction of effect is positive or negative, and examine possible sources of variation, or heterogeneity among studies. The research question was: "*What is the effect of gamified class in K-12 and university contexts on student academic performance in information science, computer and engineering education compared to traditional class?*"

3. Methods

3.1 Databases Searched

The study applied a systematic review at the beginning to select eligible papers for data extraction in meta-analysis. The process for paper selection was guided by the preferred-reporting of items for systematic reviews and meta-analyses (PRISMA) statement (Moher et al., 2015). We examined the online databases of *ACM Digital Library*, *EBSCO*, *IEEE Xplore Digital Library*, *INSPEC*, *ProQuest*, *Scopus*, *Web of Science* as they have been recommended in the fields of information, computer and engineering searching (Cavacini, 2015; Radjenović, Heričko, Torkar, & Živkovič, 2013). We included all published works in conference proceedings, journals, and full-text dissertations. The search string used in the literature search was: *gamif* AND (education OR class OR course OR learning OR performance OR behavior OR outcomes OR evaluation OR impacts OR effects OR influence)*.

3.2 Inclusion Criteria

To be eligible for inclusion in the present meta-analysis, studies had to meet the following criteria:

- (a) The focus of the study was on the field of information, computer and engineering education;
- (b) The study compared student learning achievement between a gamified class versus a non-gamified class;

- (c) The outcome variable must be based on an objective assessment such as final exams, instead of student self-reported perceptions;
- (d) The study clearly described the game design elements used;
- (e) The study reported sufficient data for calculating effect size (e.g., sample size, mean, standard deviation);
- (f) The study had to be written in English, although no restrictions were imposed on the geographical locations in which the study was conducted.

3.3 Screening

A total of 13 out of 1365 papers met the inclusion criteria. Our initial search using the aforementioned databases yielded 1365 papers (Figure 1). 167 were found to be duplicate papers due to our use of multiple academic databases. Subsequent screening of paper title and abstract removed 1079 papers due to little relevance to the research topic. The remaining 119 papers were downloaded, and their full texts were carefully read. But 106 did not meet our meta-analysis inclusion criteria (see Figure 1). Finally, a total of 13 empirical studies employing the between-subject research design were included. Two coders independently examined the excluded papers and extracted the information from 13 articles. Inter-coder reliability was 92%. The discrepancies were fully discussed and then resolved.

3.4 Effect Size Calculation

Key information such as the number of game elements used, the types of game elements, statistical data (e.g., sample size, mean, standard deviation), the setting of the interventions was extracted from each of the 13 papers. Two coders extracted the information and resolved any discrepancies of the coding. Effect sizes were computed using the Comprehensive Meta-Analysis software. All reported p-values were two-tailed. I^2 test was present for the test of heterogeneity. We used the random-effects model to compute the effect sizes since this model can account for variation in different study implementations (Raudenbush, 2009). Hedges's g , which is the adjusted standardized mean difference between two groups based on the pooled standard deviations, was used to report the effect sizes because it is particularly useful for the meta-analysis of studies with varying sample sizes (Korpershoek, Harms, de Boer, van Kuijk, & Doolaard, 2016). One effect size was calculated for each study to meet the assumption of the independence of the effect sizes based on independent samples of students. If a study reported multiple assessments of a single course subject, we selected the most summative assessment, as suggested by Freeman et al. (2014). For example, we chose final examinations over other assessments. If a study reported multiple assessments (e.g., test 1, test 2, ...) Without the most summative assessment or overall result from the same set of participants, we computed a single combined mean effect size using the Comprehensive Meta-Analysis software. To identify the possible source of variance on effect sizes, moderator analyses were performed on several variables (see Table 1).

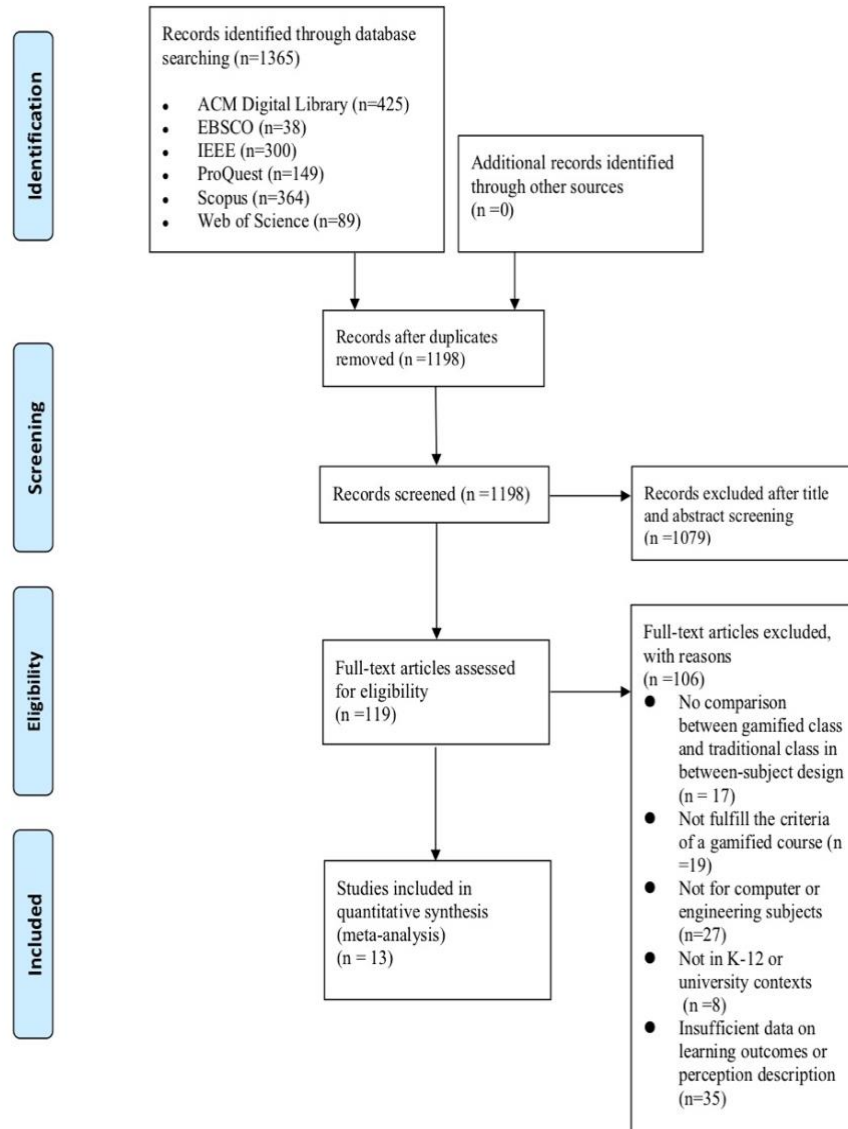


Figure 1. PRISMA flowchart of articles screening

4. Results

13 independent studies were selected for this meta-analysis with a total of 934 subjects in the gamified groups and 673 subjects in the non-gamified groups. An overall significant effect size (Hedges's $g = 0.36$, $CI = 0.006 - 0.714$, 95% confidence interval, $p = 0.046$) in favor of gamification under random-effects model (see Figure 2) was found. This suggests that the use of gamification results in higher learning gains compared with the non-gamified courses within the contexts of information science, computer and engineering education. A significant Q statistic ($p < 0.001$) indicated the presence of heterogeneity ($I^2=90.686\%$).

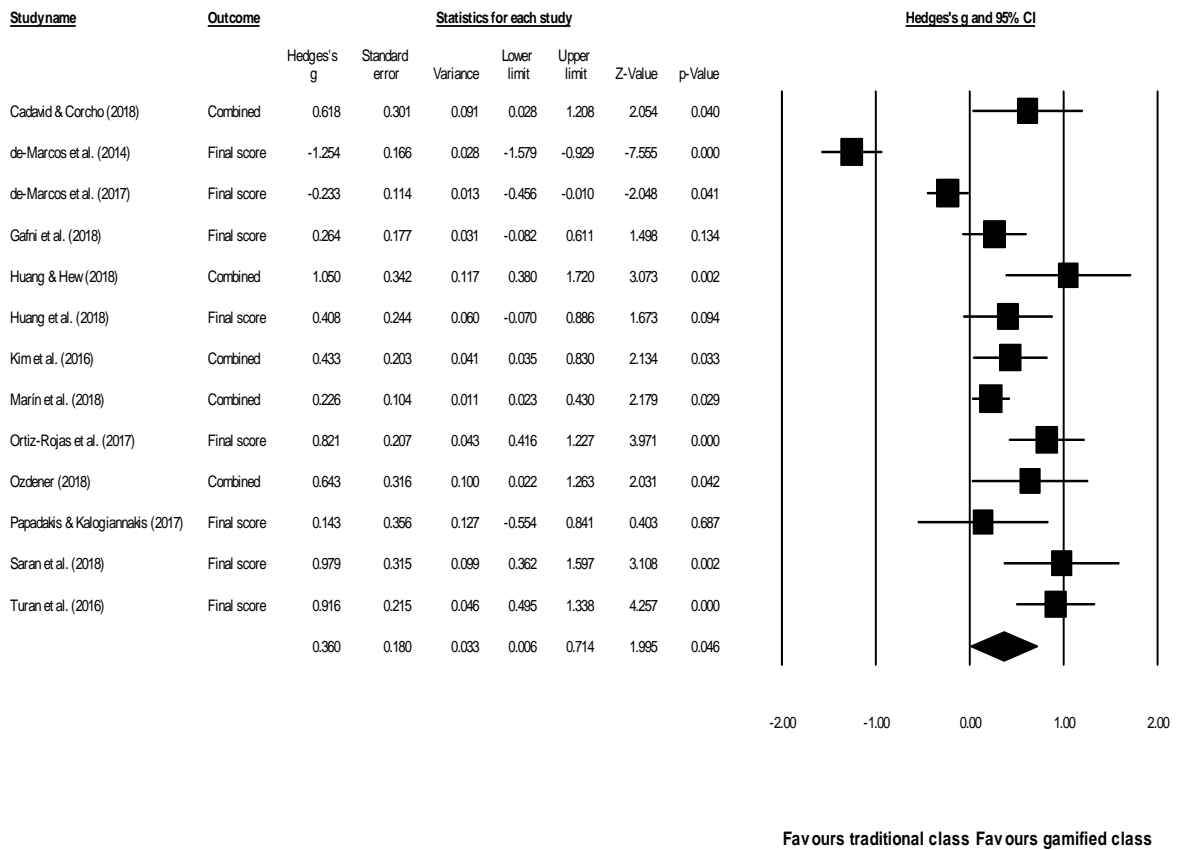


Figure 2. Forest plot of effect sizes (Hedges's g) using the random-effects model

4.1 Publication Bias

Publication bias refers to situations when authors deliberately publish significant-only results. To determine whether the results of our meta-analysis were affected by publication bias, we conducted the following tests: funnel plot, Begg and Mazumdar rank correlation, Egger's regression, fail-safe N test, and Duval and Tweedies' trim and fill. The funnel plot revealed a slightly asymmetrical position between negative and positive figures in Figure 3. However, Begg and Mazumdar rank correlation (Kendall's Tau with continuity correction) results of 0.269 (two-tailed $p = 0.2$), and Egger's regression intercept value of 3.64 (two-tailed $p = 0.121$) suggested that the overall mean effect size of 0.360 was not inflated by publication bias. Fail-safe N test showed 61 missing studies were required to bring the p -value over the alpha level (0.05). Given that we have used a broad search string, a large number of databases, as well as including both journal and conference publications, we believe that 61 missing studies would be an unreasonably large number of undetected studies with zero effect to bring the reported effect size of 0.360 to a statistically insignificant value. Due to these reasons, we believe that the overall mean effect size was not affected by publication bias.

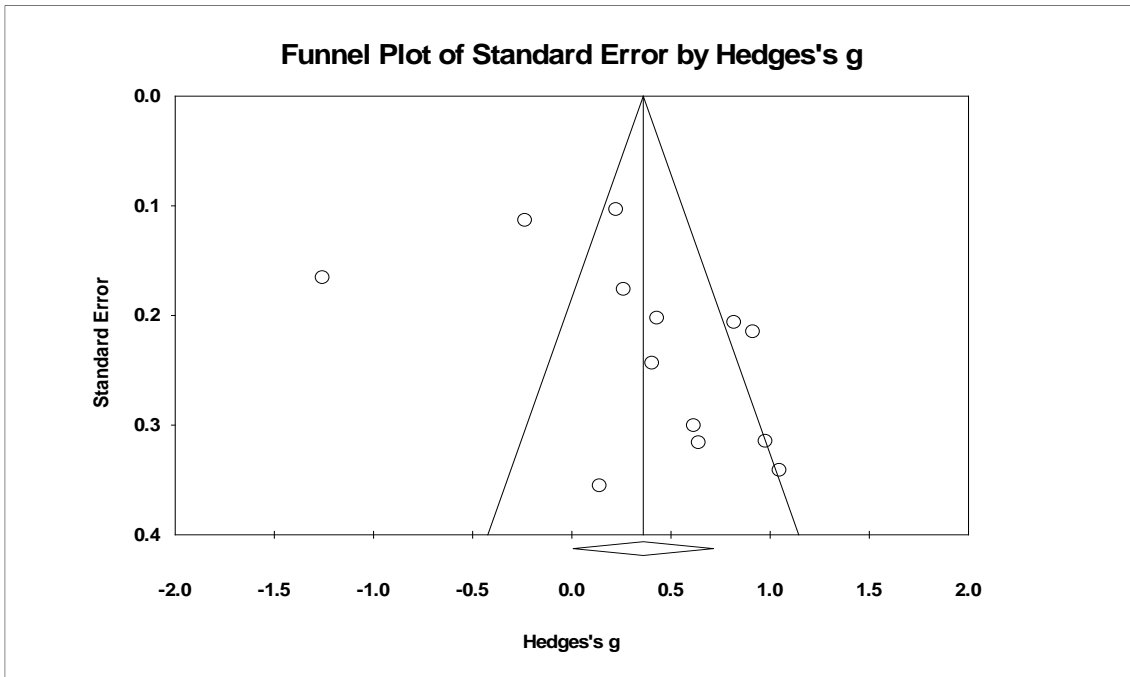


Figure 3. Funnel plot of standard error by Hedges's g under the random-effects model

4.2 Moderator Analyses

We performed several moderator analyses to explore the causal effects of student learning performance in gamified class. Three main categories of moderators were examined (Table 1): (a) controls of game design elements (the number and types of game elements used), (b) research design (research design type, student initial equivalence, and instructor equivalence), and (c) intervention condition (school setting, intervention duration and origin of study). The combination of badges, leaderboard/rank, level/unlock and points ($n=4$) was most commonly employed in the previous studies, followed by the combination of badges, leaderboard/rank, points ($n=2$). The origin of studies is listed in Table 1 and Figure 4. Spain, Hong Kong and Turkey showed high interests in gamification practices with regards to student learning performance measurement in information, computer and engineering science education.

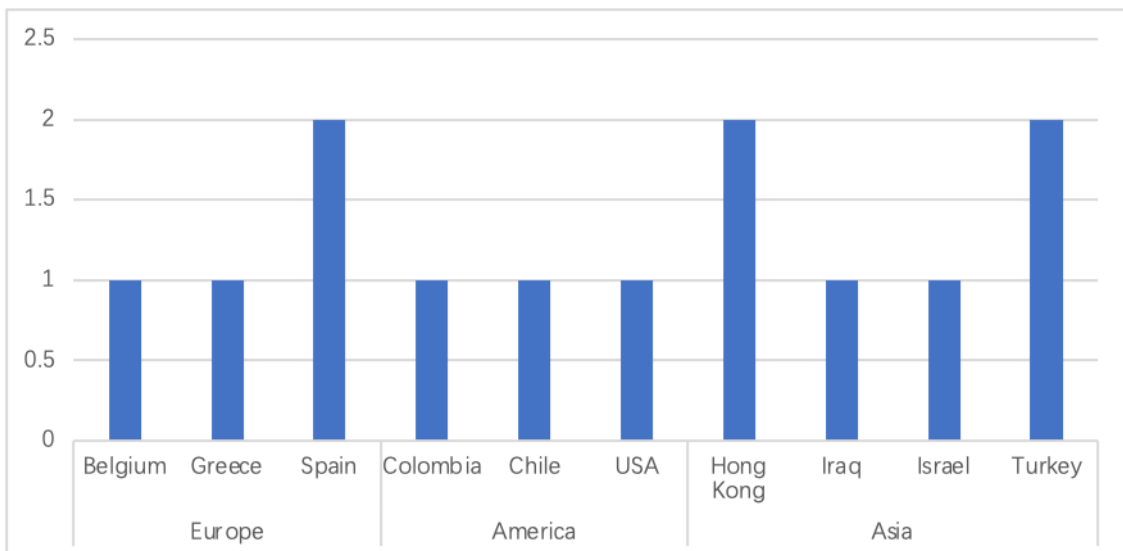


Figure 4. Origin of studies (continents and regions)

Table 1

Categories of moderator variables with numbers of studies

Moderator variables	Categories	Studies
Controls of game design elements		
Number of game design elements	1	1
	2	2
	3	4
	4	5
	6	1
	Type of game design elements	Badges + leaderboard/rank + level/unlock + points
Badges + leaderboard/rank + points		2
Badges + points + progress bar		1
Badges + leaderboard/rank + level/unlock + progress bar		1
Badges + leaderboard/rank + level/unlock		1
Badges		1
Badges + leaderboard/rank		1
Level/rank + points		1
Avatar + badges + leaderboard/rank + level/unlock + points + word notification		1
Research design quality		
Research design type	Quasi-experiment	9
	Randomization	4
Student initial equivalence (<i>i.e., whether students were indistinguishable before study</i>)	No statistical difference (pre-test)	8
	Randomized assignment	3
	No data reported	2
Instructor equivalence (<i>i.e., whether the same instructor taught the different groups</i>)	Different instructors	2
	Identical instructor	5
	No data reported	6
Intervention condition		
School setting	Elementary school	1
	High school	1
	High school + undergraduate + postgraduate	1
	Undergraduate	9
	Postgraduate	1
Intervention duration	A (one quarter of a term)	3
	B (two-quarters of a term)	3
	C (three-quarters of a term)	4
	D (more than one term)	1
	No data	2
Origin of study (continent)	America	3
	Asia	6
	Europe	4

Note: one academic term assumed to be 16 weeks (USA university standard)

We found no evidence of heterogeneity between studies that used more game elements and studies that used lesser game elements ($Q = 5.107$, $df = 4$, $p = 0.276$) (Table 2). However, results suggest a significant variation in terms of the types of game elements used ($Q = 16.007$, $df = 8$, $p = 0.042$). Effect sizes appear to be larger when badges + leaderboard/rank (Hedges's $g = 0.916$) or badges alone (Hedges's $g = 0.821$) were employed (Table 3). We observed an unexpectedly very large negative effect size (-1.254) in a single study that employed the combination of badges, leaderboard/rank,

level/unlock rendered a negative effect size (De-Marcos, Domínguez, Saenz-de-Navarrete, & Pagés, 2014), where the authors claimed the experimental group overemphasized skill acquisition in practical activities, leading to poorer scores on knowledge acquisition compared with traditional class. These results, nevertheless, have to be viewed with caution due to the very small number of studies involved.

Heterogeneity analyses indicated no significant variation when comparing (a) studies with different research design such as quasi-experiment, or true experiment ($Q = 0.02$, $df = 1$, $p = 0.889$); (b) studies that reported initial student equivalence, or not ($Q = 1.647$, $df = 2$, $p = 0.439$); and (c) studies with same or different instructors ($Q = 1.108$, $df = 2$, $p = 0.575$). There was also no evidence of heterogeneity when comparing (a) studies with different school setting ($Q = 1.892$, $df = 4$, $p = 0.756$); (b) studies with different intervention duration ($Q = 6.054$, $df = 4$, $p = 0.195$); and (c) studies conducted in different regions ($Q = 5.005$, $df = 2$, $p = 0.082$). Table 2 summarizes the results of various moderator analyses.

Table 2
Results of Q-test for heterogeneity for three main moderators

Moderator	<i>Q</i>	<i>Df</i>	<i>P-value</i>
Controls of game design elements			
Number of game design elements	5.107	4	0.276
Type of game design elements	16.007	8	0.042*
Research design quality			
Research design type	0.02	1	0.889
Student initial equivalence	1.647	2	0.439
Instructor equivalence	1.108	2	0.575
Intervention condition			
School setting	1.892	4	0.756
Intervention duration	6.054	4	0.195
Origin of study (continent)	5.005	2	0.082

* $p < 0.05$

Table 3
Effect sizes of type of game design elements

Type of game design elements	95% CI				
	<i>N</i>	Hedges's <i>g</i>	<i>SE</i>	<i>LL</i>	<i>UL</i>
Badges + leaderboard/rank + level/unlock + points	4	0.665	0.258	0.16	1.171
Badges + leaderboard/rank + points	2	0.129	0.348	-0.554	0.811
Badges + points + progress bar	1	0.264	0.474	-0.665	1.194
Badges + leaderboard/rank + level/unlock + progress bar	1	0.408	0.503	-0.578	1.394
Badges + leaderboard/rank + level/unlock	1	-1.254	0.47	-2.176	-0.332
Avatar + badges + leaderboard/rank + level/unlock + points + word notification	1	0.433	0.485	-0.517	1.383
Badges	1	0.821	0.486	-0.132	1.774
Badges + leaderboard/rank	1	0.916	0.49	-0.044	1.877
Level/rank + points	1	0.143	0.566	-0.966	1.252

Note: *n* number of studies, *SE* standard error, 95% CI 95% confidence interval, *LL* lower limit, *UL* upper limit

5. Conclusion and Discussion

This review provides a useful snapshot of the current quantitative studies on gamification with the contexts of information science, computer and engineering education. The results of this meta-analysis suggest that gamification can increase student learning performance compared to non-gamified courses. Nevertheless, it is important to note that the overall effect size of 0.360 reported here is considered a small effect (Cohen, 1988). In other words, gamification has a significant positive effect on student achievement, though the effect size is small under a random-effects model.

To our knowledge, this is the first meta-analysis conducted on gamification and its impact on student learning performance as measured using objective instruments such as student exam scores. The search string and databases were broad, thus allowing us to capture as many empirical studies as possible. The publication bias was calculated to avoid that only positive result studies were included, and no publication bias was found. Our comprehensive and careful selection of studies revealed only 13 such articles. Most of studies applied *badges + leaderboard/rank + levels + points* in combination. *Badges* and *badges + leaderboard/rank* remain top two for effect size figure, but this finding should be interpreted with caution due to the small frequency. Based on self-determination theory, humans possess three innate psychological needs to facilitate motivation to engage or not engage in an activity – autonomy, relatedness, and competence (Ryan & Deci, 2000). As one component of intrinsic motivation, the need of competence is more likely to be satisfied when students are given the autonomy to adjust difficulty levels. Therefore, *badges* and *leaderboards* can fulfil the most positive effects when students can obtain different badges/ranks when corresponding difficulty levels of learning tasks are accomplished. For game elements selection, *progress bar*, *leaderboards* and *badges* are often used to gain instant feedback, achievements and self-recognition. Students can make choices on learning tasks when *levels* and *avatars* are applied. *Points*, *badges* and *progress bar* can support goal-setting commitment.

We conclude by highlighting some follow-up research directions. One limitation of this review is that the search was limited to English and studies conducted within the contexts of information science, computer and engineering education. One follow-up research direction is to expand the search to include studies carried out in other disciplines. We noted that more high quality, rigorously designed experimental studies are needed in the field of gamification within the K-12 and university education settings. Carrying out such studies can help us better establish the causal effect of gamification on student learning. To date, most of the previous studies had only short intervention duration of less than one academic term (16 weeks). Short-term studies run the risk of novelty effect where the participants were interested to try out the intervention because it is something new and exciting. Longer-term longitudinal studies are needed to examine whether the effects of gamification on student learning hold over time.

6. References (* refers to studies included for meta-analysis)

- *Cadavid, J. n. M., & Corcho, A. s. F. P. (2018). Competitive programming and gamification as strategy to engage students in computer science courses. *Revista ESPACIOS*, 39(35), 11.
- Cavacini, A. J. S. (2015). What is the best database for computer science journal articles? , 102(3), 2059-2071. doi:10.1007/s11192-014-1506-1
- Cohen, J. (1988). *Statistical power ANALYSIS for the Behavioral sciences* (Vol. 2nd). Hillsdale, NJ: Lawrence Erlbaum Associates.
- *De-Marcos, L., Domínguez, A., Saenz-de-Navarrete, J., & Pagés, C. (2014). An empirical study comparing gamification and social networking on e-learning. *Computers & Education*, 75, 82-91. doi:10.1016/j.compedu.2014.01.012
- *De-Marcos, L., Garcia-Cabot, A., & Garcia-Lopez, E. (2017). Towards the Social Gamification of e-Learning: A Practical Experiment. *International Journal of Engineering Education*, 33(1), 66-73.
- Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). *From game design elements to gamefulness: defining "gamification"*. Paper presented at the Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, Tampere, Finland.
- Dicheva, D., Dichev, C., Agre, G., & Angelova, G. (2015). Gamification in Education: A Systematic Mapping Study. *Educational Technology & Society*, 18(3), 75-88.

- *Eunsik, K., Rothrock, L., & Freivalds, A. (2016, 12-15 Oct. 2016). *The effects of Gamification on engineering lab activities*. Paper presented at the 2016 IEEE Frontiers in Education Conference (FIE).
- Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(23), 8410-8415. doi:10.1073/pnas.1319030111
- *Gafni, R., Achituv, D. B., Eidelman, S., & Chatsky, T. (2018). The effects of gamification elements in e-learning platforms. *Online Journal of Applied Knowledge Management*, *6*(2), 37-53.
- Gari, M. R. N., Walia, G. S., & Radermacher, A. D. (2018). *Gamification in Computer Science Education: a Systematic Literature Review*. Paper presented at the ASEE Annual Conference & Exposition.
- *Huang, B., & Hew, K. F. (2018). Implementing a theory-driven gamification model in higher education flipped courses: Effects on out-of-class activity completion and quality of artifacts. *Computers & Education*, *125*, 254-272. doi:10.1016/j.compedu.2018.06.018
- *Huang, B., Hew, K. F., & Lo, C. K. (2018). Investigating the effects of gamification-enhanced flipped learning on undergraduate students' behavioral and cognitive engagement. *Interactive Learning Environments*, 1-21. doi:10.1080/10494820.2018.1495653
- Koivisto, J., & Hamari, J. (2019). The rise of motivational information systems: A review of gamification research. *International Journal of Information Management*, *45*, 191. doi:10.1016/j.ijinfomgt.2018.10.013
- Korpershoek, H., Harms, T., de Boer, H., van Kuijk, M., & Doolaard, S. (2016). A Meta-Analysis of the Effects of Classroom Management Strategies and Classroom Management Programs on Students' Academic, Behavioral, Emotional, and Motivational Outcomes. *86*(3), 643-680. doi:10.3102/0034654315626799
- *Marín, B., Frez, J., Cruz-Lemus, J., & Genero, M. (2019). An Empirical Investigation on the Benefits of Gamification in Programming Courses. *ACM Transactions on Computing Education (TOCE)*, *19*(1), 1-22. doi:10.1145/3231709
- Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., . . . Reviews, P.-P. G. J. S. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *4*(1), 1. doi:10.1186/2046-4053-4-1
- Narasareddygari, M. R., Walia, G., & Radermacher, A. (2018). *Gamification in Computer Science Education: a Systematic Literature Re-view Gamification in Computer Science Education: A Systematic Literature Review*.
- *Ortiz-Rojas, M., Chiluiza, K., & Valcke, M. (2017). *Gamification in computer programming: Effects on learning, engagement, self-efficacy and intrinsic motivation*. Paper presented at the Proceedings of the 11th European Conference on Games Based Learning, ECGBL 2017.
- *Özdener, N. (2018). Gamification for enhancing Web 2.0 based educational activities: The case of pre-service grade school teachers using educational Wiki pages. *Telematics and Informatics*, *35*(3), 564-578. doi:10.1016/j.tele.2017.04.003
- *Papadakis, S., & Kalogiannakis, M. (2017). *Using Gamification for Supporting an Introductory Programming Course. The Case of ClassCraft in a Secondary Education Classroom*.
- Radjenović, D., Heričko, M., Torkar, R., & Živković, A. (2013). Software fault prediction metrics: A systematic literature review. *Information and Software Technology*, *55*(8), 1397-1418. doi:https://doi.org/10.1016/j.infsof.2013.02.009
- Raudenbush, S. W. (2009). Analyzing effect sizes: Random-effects models. In *The handbook of research synthesis and meta-analysis, 2nd ed.* (pp. 295-315). New York, NY, US: Russell Sage Foundation.
- Ryan, R. M., & Deci, E. L. (2000). Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *American Psychologist*, *55*(1), 68-78. doi:10.1037/0003-066X.55.1.68
- Sailer, M., Hense, J. U., Mayr, S. K., & Mandl, H. (2017). How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. *Computers in Human Behavior*, *69*, 371-380. doi:10.1016/j.chb.2016.12.033
- *Saran, M., Al-Magsoosi, D., & Mohammed, D. (2018). *Gamification in e-learning: The Effect on Student Performance*.
- Souza, Mauricio R. de A., Veado, L., Moreira, R. T., Figueiredo, E., & Costa, H. (2018). A systematic mapping study on game-related methods for software engineering education. *Information and Software Technology*, *95*, 201-218. doi:https://doi.org/10.1016/j.infsof.2017.09.014
- Subhash, S., & Cudney, E. A. (2018). Gamified learning in higher education: A systematic review of the literature. *Computers in Human Behavior*, *87*, 192-206. doi:10.1016/j.chb.2018.05.028
- *Turan, Z., Avinc, Z., Kara, K., & Goktas, Y. (2016). Gamification and Education: Achievements, Cognitive Loads, and Views of Students. *International Journal of Emerging Technologies in Learning*, *11*(7), 64-69. doi:10.3991/ijet.v11i07.5455