Does Gamification Significantly Improve Mathematics Achievement among Junior High School Learners?: Evidence from a Meta-Analysis

John Raubensun D. Hipulan[1] and Angel Constantine Bajana[2], MAED

[1] Senior High School Teacher II /
F. Bustamante National High School, Davao City
johnraubensun@gmail.com
[2] University of the Immaculate Conception

ABSTRACT

National and international standards have measured mathematics achievement among junior high schools and the results were very low. This research employed a systematic review and meta-analysis to determine if gamification significantly improves the mathematics achievement of Junior High School students compared to traditional instruction. In gathering data, the search approach in various online databases produced 157 studies, where only 15 studies, carefully selected based on their titles and abstracts, fit the inclusion criteria which the respondents were in junior high school, published between 2010 to 2022, and using a quasi-experimental design. These studies were systematically reviewed in a participant-intervention-comparison-outcome table. Further, Jadad's scoring was utilized by professional experts to ensure high-quality research output in which four studies qualified for meta-analysis. It was determined that the overall Standardized Mean Difference (SMD) of the four studies was statistically insignificant. However, the heterogeneity table revealed an accepted level of heterogeneity. The funnel plot showed no evidence of publication bias.

KEYWORDS: *Education, gamification, quasi-experimental, systematic review, mathematics, Philippines*

INTRODUCTION

Mathematics proficiency is the intertwining five strands of development of conceptual understanding, procedural fluency, strategic competence, adaptive reasoning, and productive disposition (Kilpatrick et al., 2001, cited in Groves 2012). With this, the Organization for Economic Cooperation and Development (OECD) carries out the Program for International Student Assessment (PISA), an examination of students worldwide which administered a computer-based assessment for almost 80 countries in math, science, and reading. The Philippines only scored 353 compared to the OECD average score of 489 (OECD, 2018). This is the second lowest mathematics literacy behind 76 participating countries. Moreover, only 1 out of 5 Filipino students attained a Level 2 proficiency in mathematics, which is 26 points behind Indonesian students in mathematics. This means that Filipino students have low skills in conceptualizing math problems, lack investigation and modeling skills, and lag in advanced mathematical thinking and reasoning. Filipino learners are at the level of answering questions involving familiar contexts, can only determine the information and execute actions in accordance with explicit instructions, and can only perform actions that are always almost obvious (OECD, 2018). As a result, Filipino students' math performance is much worse than students in many other nations.

Further, the educational system in the country has widely been adapting to different learning tools other than conventional learning to improve its national performance and even international standards. One of the learning tools that has been used is gamification tools. The present generation is using gamification as a learning tool more frequently. As a result, several studies have been done nationally and internationally to determine whether gamification is helpful in many subject areas, including language, math, and science. This new learning system has been used in over 237,000

research that has been found in Google Scholar, over 28,819 studies in Bielefeld Academic Search Engine (BASE), about 34,400 articles in Semantic Scholar, and over 58,998 studies in ProQuest. There are almost 26,300 publications on gamification and mathematics education in Google Scholar, 709 entries in BASE, 294 articles in Semantic Scholar, and 9,535 articles in ProQuest. With the number of gamification-related publications being published, which are propagating and building up the research databases over time, these numbers are consistently rising. Most studies from the published research are experimental and frequently applied to students with varied levels of education. These studies often compare the effectiveness of gamification with conventional teaching methods using independent variables of academic performance, engagements, and motivation or to target skill development.

Furthermore, many studies show a more positive impact of gamification in mathematics than conventional instruction. There is research examining the efficiency of gamification as a pedagogy in mathematics. Text, images, sound, and animation from multimedia sources serve an increasing student enthusiasm for studying through interactive math instruction (Udjaja et al., 2018). In addition, gamified STEM activities help the students learn and practice the mathematics material, leading them to consider the activities fun, thrilling, and competitive, indicating that they have high levels of intrinsic motivation (Asigigan & Samur, 2021). Students are more likely to have higher learning expectations if they have higher levels of internal motivation and a favorable attitude toward gamification. In the study conducted by Sutopo (2017), students can solve mathematics through mobile games. This supports the idea that online games are an essential tool to support learning. With this, it is evident that gamification as a learning tool can be successful in managing students' motivations in learning activities. Thus, it can help ensure that kids achieve success in both their academic and social lives (Lee & Hammer, 2016).

Additionally, due to the spread of the COVID-19 pandemic, online game-based learning has been incorporated to increase engagement among students during synchronous classes. A study conducted in one of the schools in the City of Biñan reveals that while students' attitudes about gamification in mathematics have not changed considerably, subject performance has improved significantly (Antonio & Tamban, 2022). Thus, students' performance in mathematics using gamification improves online distance learning. In support of the idea, Libradilla et al. (2015) found learning using gamification to be more engaging, interesting, and thrilling because of gaming aspects, including appealing color, images, music, and time element.

Moreover, online technology is an essential tool for learning in this generation. Amidst the relevantly low level of readiness in technology education, some schools in Mindanao, Philippines slowly embrace digital-based learning to provide quality education. Specifically, the study conducted by Boldadora (2018) shows that in one of the schools located in Davao City, there is a significant increase in mathematics achievement using digital game-based learning. One of the most used game-based platforms in the region, namely, Kahoot! is considered for it is accessible and free of charge. According to Baguio et al. (2021), a group of Grade 8 students in one of the schools in Northern Mindanao that uses the platform showed improved test scores in Geometry, especially in formative assessments. It is evident that motivation is relevantly high. With this, it is concluded that gamification in math learning influences the achievement and motivation of online classes.

For years, numerous studies have demonstrated how effectively gamification has raised students' mathematics achievement levels. This learning tool has been used on many levels as integrated into education. However, mathematics achievement in Junior High School is still very low, contrary to the results of the studies about gamification. The goal of this study is to give researchers and practitioners a better understanding of how well gamification works as a method for teaching

math. Therefore, this study aims to determine if gamification helps improve mathematics achievement among junior high school learners.

METHODS

Research Design. The study used systematic review and meta-analysis to determine whether using gamification as a learning tool would significantly enhance mathematical proficiency. The most vital type of evidence for combining results from well-conducted studies from various sources with comparable objectives is provided by this method. The systematic review was utilized to examine the data and assess whether gamification of learning would significantly increase student performance in mathematics. Therefore, this study focuses on the significant difference in the use of gamification in mathematics achievement between students in conventional learning.

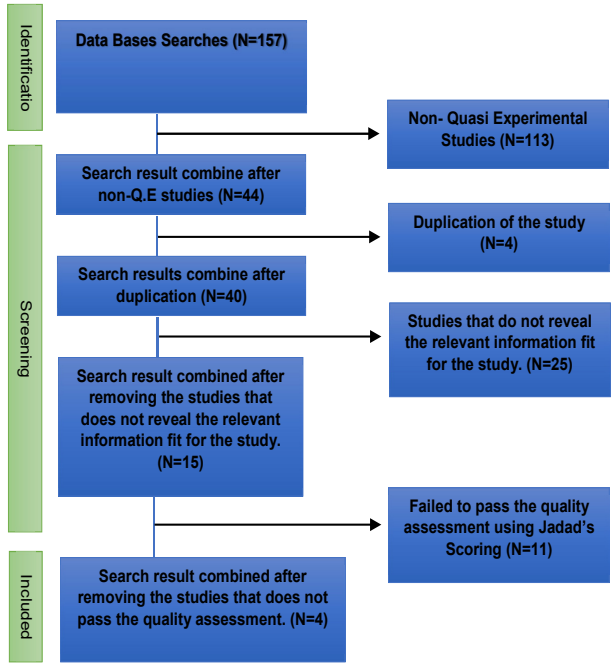
Place of Study. The research was conducted at the University of the Immaculate Conception Davao City, library hub, internet cafes, workplaces, and public places with available high-speed Wi-Fi connection. The researcher utilized electronic databases and the available resources in the ProQuest World Wide Web (e.g., Google Scholar, ProQuest, ERIC, PsycINFO, and PubMed) for more references to limit the time used for looking through the theses and dissertations. All the studies and the inclusion and exclusion criteria to be implemented in selecting the studies are set.

Data Sources and Selection. The purpose of this study is to accumulate an understanding of the status quo of the effectiveness of gamification used as a learning tool for mathematics proficiency at the high school level. The electronic databases used for this study are ProQuest Educational Research Journal, Springer Link, Open Access Journal, Education Resources Information Center (ERIC), Google Scholar, Education Journal, etc. Abstracts were deemed eligible for complete manuscript data extraction if they met all of the following requirements: they had to be quasi-experimental studies on the use of gamification to improve student achievement in mathematics; the authors had to have used an outcome variable to assess students' academic performance and achievement; the studies focus on the Junior High School students that utilize the use of the test for math achievement; and the studies had to have specifically measured the effectiveness of gamification; the independent variables were quantified and measured using appropriate assessment tools, and the dependent variable was also quantified or measured using grades or other assessment procedures that the authors used. The studies were published from 2010 to 2022 and included students in grades 7 to 10, which reflected the mean difference, p-value, t-value, and sample size of the data. Exclusion criteria were also established. These were non-quasi-experimental research, studies that were published twice and failed to provide the necessary data for the study are one of the exclusion criteria. Three experts assessed the quality of each paper to ensure quality outputs using Jadad's Scoring sheet. All papers that meet the inclusion criteria and pass the quality assessment were thoroughly examined in the current evaluation. The statistical information for each chosen study, including the title, authors, total sample size (n), p-value, t-value, and degree of significance, were organized using a coding sheet.

Data Synthesis and Analysis. The groups' means and SDs, and effect sizes, such as SMDs, Q-test, I-square test, forest plot analysis, and funnel plot analysis, were employed in the meta-analysis phase of this study to combine the quantitative findings from the numerous reputable research investigations. Moreover, funnel plot, fail-safe N, and trim and fill methods were utilized for the publication bias of the paper.

RESULTS AND DISCUSSION

There were four studies included with multiple authors. Three (3) studies were authored by one (1) person, and one (1) study was authored by three (3) different individuals. Additionally, these studies were conducted in North America and Africa Regions with varying mathematics education.



Thea data in Table 1 shows the features of the included studies using the PICO (Participant, Intervention, Comparator, and Outcome) table. These utilized quasi-experimental designs with varying sample sizes and participants. The most significant number of participants was 220, and the least number of participants was 71. The probability values were >0.05 for Study 1, <0.0005 for Study 2, >0.05 for Study 3, and 0.009 for Study 4.

Table 1. Participant, Intervention, Comparator, and Outcome (PICO) Profile of the Included Studies

Title of the Study	Participants	Intervention	Comparator	Outcome
1. Mario Math with Millennials: The Impact of Playing the Nintendo DS on Student Achievement Author: Adam Gelman (2010) Place: Denver, USA **Duration 9 weeks**	Seventh Graders Experimental- 35 Control- 36	**Mario Math: Brain Age 2 Nintendo DS**	**Traditional Teaching Method**	**Group** / **EG** / **CG** Mean Post-Test: 213.86 / 218.14 **MD**: -4.28 It is not supported that playing Brain Age 2 on the Nintendo DS has an impact on math aptitude or skill as determined by the MAP test.

<table>
<tr>
<td>2. The Effects of Digital Games on Middle School Students' Mathematical Achievement
Author: Pilar Lisa Starkey (2013)
Place: Africa
Duration: 9 weeks</td>
<td>Middle School Students: 6th, 7th, and 8th Graders

Experimental- 84
Control- 84</td>
<td>Digital Games: Lure of the Labyrinth</td>
<td>Traditional Method</td>
<td>
<table>
<tr><td>Group</td><td>EG</td><td>CG</td></tr>
<tr><td>Mean Post-Test</td><td>82.60</td><td>80.61</td></tr>
<tr><td>MD</td><td colspan="2">1.99</td></tr>
</table>
According to the study's findings, the treatment group, who played Labyrinth in addition to attending traditional classes, scored better in mathematics than the control group, which only attended traditional classes.
</td>
</tr>
<tr>
<td>3. Effect Of Digital Game-Based Learning on Ninth Grade Students' Mathematics Achievement
Author: DIXIE K. SWEARINGEN (2011)
Place: Oklahoma, USA
Duration: 7 weeks</td>
<td>Ninth Grades

Experimental- 127
Control- 93</td>
<td>Interdisciplinary massive multiplayer online game (MMOG)</td>
<td>Traditional without Intervention</td>
<td>
<table>
<tr><td>Group</td><td>EG</td><td>CG</td></tr>
<tr><td>Mean Post-Test</td><td>47.13</td><td>47.20</td></tr>
<tr><td>MD</td><td colspan="2">-0.07</td></tr>
</table>
Between the treatment and control groups, the mean post-test results did not show any statistically significant differences.
</td>
</tr>
<tr>
<td>4. The effects of modern mathematics computer games on mathematics achievement and class motivation

Authors: Mansureh</td>
<td>High School students: Grades 9 and 10

Experimental – 117
Control – 76</td>
<td>DimentionM</td>
<td>Conventional Instruction</td>
<td>
<table>
<tr><td>Group</td><td>EG</td><td>CG</td></tr>
<tr><td>Mean Post-Test</td><td>45.71</td><td>32.00</td></tr>
<tr><td>MD</td><td colspan="2">13.71</td></tr>
</table>
In the context of a public high school, the DimensionM
</td>
</tr>
</table>

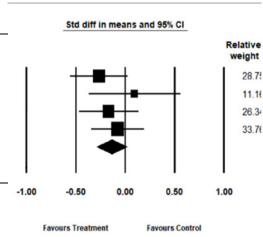
Kebritchi , Atsusi Hirumi, Haiyan Bai (2010) Place: USA Duration: 18 weeks				games had a very favorable impact on pupils' mathematical achievement. On the district-wide math benchmark exam, pupils who played computer games performed noticeably better than those who did not.

Table 2 shows the meta-analysis of the included studies with its classification of standardized mean difference, standard error, variance, confidence interval, z-value, and p-value. This table illustrates the effectiveness of gamification in math achievement.

Table 2. Forest Plot

Meta-Analysis of the Effectiveness of Gamification on in Math Achievement

Study Name	Statistics for each study						
	Std diff in means	**Standard error**	**Variance**	**Lower Limit**	**Upper Limit**	**Z-value**	**p-value**
Kebritchi et al. (2010)	-0.268	0.148	0.022	-0.558	0.022	-1.814	0.070
Gelman (2010)	0.093	0.238	0.056	-0.373	0.558	0.391	0.696
Starkey (2013)	-0.168	0.155	0.024	-0.471	0.135	-1.088	0.277
Swearingen (2011)	-0.078	0.137	0.019	-0.345	0.190	-0.569	0.570
Pooled	-0.137	0.079	0.006	-0.293	0.018	-1.731	0.084



Forest Plot.

It displays the forest plot for the standardized mean differences (SMDs) in the measurement outcomes between the experimental and control groups of the quasi-experimental studies assessing the effectiveness of gamification in mathematics achievement. The figure shows the relevant information, such as the effect size, standard error, and level of significance. Below are the discussions of the effectiveness of gamification in mathematics achievement:

Effect Size and the Null Effect Vertical Line. The four effect sizes in the meta-analysis using the standardized mean difference are shown in Table 2, ranging from -0.268 to 0.093, with 75%

of the point estimate being negative. It means that three out of 4 studies showed favorability to gamification in mathematics achievement. These are the studies of Kebritchi et al. (2010), Starkey (2013), and Swearingen (2011), who have SMDs of -0.268, -0.168, and -0.078, respectively. On the other hand, the study of Gelman (2010) favors the use of the traditional method since it has an SMD of 0.093.

The four included studies showed different values of SMD. The study by Kebritchi et al. (2010), Gelman (2010), Starkey (2013), and Swearingen (2011) revealed that it crosses the line of no effect through 0.00, denoting that the result is not significant; in other words, the p-value is greater than 0.05. Overall, the meta-analysis showed a pooled value of -0.137 and a p-value of 0.084. Further, the junior high school students have no significant improvement in the post-test results. This means that gamification has no significant difference compared to traditional methods. Thus, gamification among junior high schools showed no significant improvement in their mathematical achievements.

SMDs and 95% Confidence Intervals of Individual Studies. Also, figure 4 reveals the SMDs of the four studies reviewed, as represented by the black boxes, with their corresponding 95% confidence intervals, represented by the "whiskers" on both sides of each black box. Thus, for example, Swearingen (2011) has a higher weight that embodies a more important black box means having a higher number of participants, establishing a "narrower" confidence interval. In contrast, Gelman (2010) had a smaller and broader whisker means a smaller sample size. Kebritchi et al. (2010) showed an SMD of -0.268 with a 95% confidence interval (CI) of -0.558 to 0.022. This included study yielded a z-value of -1.814 with a large standard error (SE) of 0.148. It has been found to be insignificant alongside its p-value of 0.070. Similarly, the study of Starkey (2013) (SMD=-0.168, CI [-0.471, 0.135], z-value = -1.088, SE = 0.155, p= 0.277), Swearingen (2011) (SMD=-0.078, CI [-0.345, 0.190], z-value = -0.569, SE = 0.137, p= 0.570), and Gelman (2010) (SMD=0.093, CI [-0.373, 0.558], z-value = 0.391, SE = 0.238, p= 0.696) revealed to be insignificant. Figure 4 shows that none of the included studies has significant results about the use of gamification in mathematics achievement.

The findings contradicted those that suggest that gamification integration has significantly improved math achievement. Yldrm and Sen's (2019) finding that gamification has a moderately positive effect on mathematics achievement is supported by Tokac et al. (2019), who found that game-based learning contributes to math achievement. However, not all studies have found a benefit to using video games for math instruction in terms of effect-size heterogeneity results.

Overall SMD. Four effect size values were analyzed using a pooled value model generated by Comprehensive Meta-analysis software. Under the pooled value, the overall SMD was found to be -0.137. The test of the null hypothesis yields a z-value of -1.731 and a corresponding p-value of 0.084. These statistics values were found to be insignificant. Further, this result indicates that the control group's outcome is slightly higher than the experimental outcome. The experimental has a higher post-test score than the control group. This means that the number of studies in this paper lacks to prove the effectiveness of gamification. Correspondingly, the 95% confidence interval is -0.293 to 0.018, indicating 95% certainty that the true SMD lies within this range of values. It is also evident from Figure 4 that the diamond at the bottom of the graph touched the line of no effect; hence, the overall results were statistically insignificant. Furthermore, the diamond is more to the left; thus, more episodes of outcome favor gamification.

SMD has been demonstrated to have several advantages over alternative effect size measures, such as the odds ratio or the raw mean difference. For instance, research by Lipsey and Wilson (2001) discovered that SMD was more resistant to violations of distributional assumptions and less susceptible to outliers than other effect sizes. Hedges and Olkin (1985) observed in another

investigation that SMD was more powerful and effective at detecting treatment effects in meta-analyses than other effect sizes.

The Cochran's Q = 1.955 with p-value = 0.582 indicates that the individual studies' SMDs statistically evaluate the same effect size regarding the overall SMD. It suggests that there are indeed genuine similarities underlying the results of the studies. Since the power of Q, in this case, is not high due to the limited number of included studies, there is no hint of heterogeneity in this study. Therefore, there is an acceptable heterogeneity in the four included studies. Moreover, as revealed in this study, this collection of studies has established an I-squared = 0.000. There is a very low value of I-squared, which means there is acceptable heterogeneity in the included studies in this meta-analysis. In this study, the CMA software generated that the true effect sizes between 4 studies were dispersed at 0.000 in terms of the scale of the effect size.

A relatively small number of studies (n=4) were included in this analysis. Two studies are adequate for a meta-analysis, according to Ryan (2016) in the Cochrane Handbook for Systematic Reviews of Intervention, assuming that they can be meaningfully pooled and that their findings are sufficiently similar. Farhad (2022) agreed that two studies are enough to conduct a meta-analysis, however, limiting the included studies could be biased and produce outliers. It was also claimed that the assessment of heterogeneity allows one to understand what is "sufficiently similar." The heterogeneity table provides sensitivity analysis between studies. Table 7 showed that there was at least a minimal amount of heterogeneity among the included studies, indicating that the significance of the results and similarity of these studies was insufficient. As a result, the small number of included studies influenced the computation parameters within the studies.

Consequently, Von Hippel (2015) pointed out that I^2 has a considerable bias when there are few studies, and that the bias is positive when the fraction of heterogeneity is low, but frequently negative when it is high. There is no evidence of the study's heterogeneity because only four (4) papers were included in the meta-analysis. Additionally, according to Huedo-Medina et al. (2006), results from I^2 should be regarded cautiously when the number of included studies is limited, specifically less than 20, due to its low statistical power. The I-squared value of 0.000 indicates a minimal number of included studies in the meta-analysis, which implies a minimum level of heterogeneity. Therefore, the result of a very low value of I-squared is associated with the few numbers of included studies in this analysis,

The use of gamification as a learning approach in mathematics insignificantly improves the mathematics achievement level of students in different math contexts. The meta-analysis involving the deliberately selected studies has established that gamification or digital game-based learning with different applications such as Mario Math: Brain Age 2 Nintendo DS, Lure of the Labyrinth, Interdisciplinary Massive Multi-Player Online Games (MMOG), and Gamified Classroom will not elevate the mathematics performance of the students. The findings of the systematic review and meta-analysis revealed an insignificant difference between instruction with gamification and those with pure traditional methods of teaching. With most studies favoring the use of gamification, it was still determined that there are not enough studies supporting the effectiveness of the use of gamification. Therefore, the integration of gamification in mathematics classrooms has no significant impact on the achievement of junior high school students.

REFERENCES

Antonio, J. M. A., & Tamban, V. E. Effectiveness of Gamification on Learner's Performance and Attitude towards Mathematics Amidst the COVID-19 Pandemic.

Asigigan, S. İ., & Samur, Y. (2021). The Effect of Gamified STEM Practices on Students' Intrinsic Motivation, Critical Thinking Disposition Levels, and Perception of Problem-Solving Skills. *International Journal of Education in Mathematics, Science and Technology*, *9*(2), 332-352.

Baguio, F., Rama, F. M. I., Rico, J. M. L., & Salazar, D. A. (2021, March). Grade 8 learner's achievement and motivation level in geometry using kahoot! as a formative assessment. In Journal of Physics: Conference Series (Vol. 1835, No. 1, p. 012014). IOP Publishing.

Boldadora, R. B.. (2018). Effectiveness of Digital Game-Based Learning on Mathematics Achievement of Grade Six Students . Tin-aw, 2(1). Retrieved from http://www.ejournals.ph/form/cite.php?id=13665

Farhad. (2022, July 18). What is the minimum number of included studies in a systematic review? Medium. https://farhadinfo.medium.com/what-is-the-minimum-number-of-included-studies-in-a-systematic-review-5ce214df6c4a

Gelman, A. (2010). Mario math with millennials: The impact of playing the Nintendo DS on student achievement. University of Denver.

Groves, S. (2012). Developing mathematical proficiency. *Journal of science and mathematics education in Southeast Asia*, *35*(2), 119-145.

Hedges, L. V., & Olkin, I. (1985). Statistical methods for meta-analysis. Academic Press.

Lipsey, M. W., & Wilson, D. B. (2001). Practical meta-analysis. Sage Publications.

Huedo-Medina, T. B., Sánchez-Meca, J., Marín-Martínez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I^2 index? Psychological Methods, 11(2), 193–206. doi:10.1037/1082-989x.11.2.193

Kebritchi, M., Hirumi, A., & Bai, H. (2010). The effects of modern mathematics computer games on mathematics achievement and class motivation. Computers & education, 55(2), 427-443.

Lee J., Hammer J. (2011). Gamification in Education: What, How, Why Bother? Academic Exchange Quarterly, 15(2). Retrieved May, 2016, from http://www.gamifyingeducation.org/files/Lee-Hammer-AEQ-2011.pdf

Libradilla, H., Teves, K., & Teves, A. (2015). Teaching effectively with use of game-based interactive mathematics. In *International Conference on Trends in Economics, Humanities and Management* (pp. 169-173).

OECD. (n.d.-a). Retrieved November 5, 2022, from https://www.oecd.org/pisa/Combined_Executive_Summaries_PISA_2018

Ryan R (2016). Cochrane Consumers and Communication Review Group. Cochrane Consumers and Communication Group: meta-analysis. http://cccrg.cochrane.org, December 2016.

Starkey, P. L. (2013). The effects of digital games on middle school students' mathematical achievement. Lehigh University.

Sutopo, H. (2017). Mobile Game Developing: Math Mobile Game Learning Model. *International Journal of Information and Electronics Engineering*, *7*(2), 62-67

Swearingen, D. K. (2011). Effect of digital game based learning on ninth grade students' mathematics achievement. The University of Oklahoma.

Tokac, U., Novak, E., & Thompson, C. G. (2019). Effects of game-based learning on students' mathematics achievement: A meta-analysis. Journal of Computer Assisted Learning, 35(3), 407-420.

Udjaja, Y., Guizot, V. S., & Chandra, N. (2018). Gamification for elementary mathematics learning in Indonesia. *International Journal of Electrical and Computer Engineering (IJECE)*, *8*(6).

Von Hippel, P. T. (2015). The heterogeneity statistic I2 can be biased in small meta-analyses. BMC Medical Research Methodology, 15(1). doi:10.1186/s12874-015-0024-z

Yıldırım, İ., & Şen, S. (2019). The effects of gamification on students' academic achievement: a meta-analysis study. Interactive Learning Environments, 1–18. doi:10.1080/10494820.2019.1636089