Received: 7 September 2022 | Revised: 27 January 2023 | Accepted: 31 March 2023

DOI: 10.1002/cae.22630

RESEARCH ARTICLE

WILEY

Using game concepts to improve programming learning: A multi-level meta-analysis

Joana M. Costa

Instituto Universitário de Lisboa (ISCTE-IUL), ISTAR, Lisboa, Portugal

Correspondence
Joana M. Costa, Instituto Universitário de Lisboa (ISCTE-IUL), ISTAR, Lisboa, Portugal.
Email: joana.martinho.costa@iscte-iul.pt

Funding information
Fundação para a Ciência e a Tecnologia, Grant/Award Numbers: UIDB/04466/2020, UIDP/04466/2020

Abstract

Gamification has been widely used in education due to its application to different contexts. The vast number of possibilities allows the creation of learning strategies that lead to students' success. Despite the high acceptance of gamification among several authors in educational contexts, there are many dispersed studies with high differences in their results. Although gamification impacts motivation, interest, and engagement, its effect on students' learning outcomes needs to be clarified. This study aims to investigate the effects of gamification on programming learning and the impact of the most used game concepts on knowledge acquisition. A multi-level meta-analysis was conducted on studies above K-12 to understand the effects of gamification in programming learning. From 15 combined effect sizes, it was analyzed the effect of points, badges, levels, avatars, leaderboards, and the number of elements used. The results showed that gamification leads to a significant increase in the results (Cohen's $d = 0.4$) but essentially due to the use of levels ($p < .05$). Using badges, points, avatars, and leaderboards showed no significant effect on programming learning, suggesting that these elements may be more beneficial to motivating or engaging the students. Gamification can be efficient in programming learning, but it is recommended the use of levels.

KEYWORDS

game concepts, gamification, multilevel meta-analysis, programming learning

1 | INTRODUCTION

The most significant challenges in programming teaching and learning have been studied and debated in the last decades, including how to deal with the lack of abstract reasoning and logical skills, low level of previous mathematical knowledge, the chosen programming language to start the learning, the teaching strategy, among others (e.g., [21, 29, 52]). Besides the efforts with promising results on knowledge acquisition (e.g., [29, 24]), research on programming teaching and learning is still looking for a solution to mitigate these problems and simultaneously engage and motivate students. One of the proposed solutions in the literature refers to the application of gamification in the process, which materializes using gamified learning activities and software (e.g., [47]).

Gamification has been described in multiple ways due to its cross-cuting applications across different contexts. The most consensual definition of gamification involves incorporating game design elements into a nongame context [13, 44]. In educational environments, it is used strategically to increase student engagement, motivation, and learning.

Using gamification in the context of education paves the existence of at least one game mechanism (e.g., points, leaderboards, progress bars, and avatars), dynamic (e.g., badges and achievements), or element (e.g., levels, immediate feedback) [57]. The many possibilities of application of game concepts, combined or individual, allow teachers to create learning strategies that lead students to success.

Sailer et al. [43] presented a set of motivation mechanisms distributed by game elements: the use of points acts as positive reinforcement, badges as a motivating factor, leaderboards for social relatedness, and the building of avatars to promote autonomy, among others. However, Sailer et al.'s [43] study does not include using game concepts to improve learning. Moreover, the effort and time to prepare a gamified lesson may not be worth it if the results are not evident [36]. On the same line, Schöbel et al. [46] conducted a bibliometric analysis of game concepts in digital learning environments. They also encourage gamification, but their results did not include articles presented in the Scopus database, only Web of Science, limiting the results. Besides, they point out the need for further research to understand the effect of game mechanisms.

1.1 | Engagement and motivation

Gamification has been widely used to motivate, engage and increase students' satisfaction levels, and may include a single or a set of gamified elements [55].

The influence of badges on motivation and engagement was researched in some studies. It is the case of Facey-Shaw, Specht, and Bartley-Bryan's [15] and Azmi et al.'s [2] studies. Both concluded that badges are positive elements of motivation, where students suggested using physical and competitive badges to increase their motivation. On the same line, Bogdanovych and Trescak [6] changed a programming course to a gamified version using star ratings, badges, and challenges with user-created content. The results showed a high increase in the interest and engagement of students.

However, not all studies present positive results. Tomaselli et al. [56] developed a study with 717 participants and concluded that contrary to the expected, points, badges, and leaderboards were not the main factor in students' engagement. The authors argue that the described distinct effects of gamified elements can be related to how they are used. So, on the one hand, they be suitable for setting goals but, on the other hand, can stress peer competition, as seen in Bai et al.'s [3] work. In this study, the authors suggest that the absolute leaderboard helps to enhance students' sense of comparison and competitiveness more than a relative leaderboard once students in different positions showed similar levels of learning performance and course engagement. In the relative class of the leaderboard, students ranked in the top third tend to perform better in learning and present higher levels of intrinsic motivation than their peers in the bottom two-thirds.

The review of Venter [59] indicates that student engagement varies across several studies from positive (most of the studies) to no impact at all (only one study), and student motivation shows a positive impact in all. However, its programming knowledge presents more heterogeneous results, varying between positive (five studies), negative (four studies), and no impact (one study). These results reinforce the need to understand how far and which gamified elements impact programming learning.

1.2 | Attitudes

Some studies have identified differences in results according to students' characteristics. Smiderle et al. [50] and Smiderle et al. [51] found that introverted students demonstrated more engagement and correct exercises using points and badges than extroverted students. Contradictory results have been found in Jia et al.'s [23] work, where extrovert students showed to be more motivated by points, ranking, and levels.

Shürmann and Quaiser-Pohl [49] conducted a quasi-experimental study for 8 weeks with 64 participants through pre-test and posttest to assess the satisfaction and frustration needs. Participants were divided into an experimental (using badges) and control group (not using badges). Male-gender people have perceived working with badges as less satisfying than working for the seminar in general. Frustration was unaffected for both genders. Qualitative data showed that autonomy, competence, and relationship needs could be better supported by highlighting the badge's value, ease of use, and social aspects. The authors consider that gender effects may stem from a male preference for competitive game elements or ease of use.

Cuervo-Cely et al. [11] did not analyze the gender. However, they developed a quasi-experimental mixed-explanatory study with 48 participants to analyze the use of computer-assisted gamification in a computer programming course. The results showed that the experimental group felt authentic desires to learn, greater self-confidence to approach learning tasks, and better expectations to achieve their learning objectives in the discipline.

Several studies have analyzed the time that students pass with a gamified tool. Lehtonen et al. [27] verified that students increased the number of solved exercises and time during exercises when they used gamification to the detriment of non-gamification. However, the number of solved exercises and time passed on the platform does not mean that learning or knowledge was acquired. For example, according to Featherstone and Habgood [16], more time can be related to cheating, as good students achieve proficient results, and bad students pass, primarily due to self-claiming tasks.

Rodrigues et al. [41] presented a longitudinal study in which they evaluated the use of gamification in programming learning in a 14-week course with 756 students. Their findings revealed that the gamification effect diminished after 4 weeks for all behavioral measures (attempts, use, and access). These findings corroborate the novelty effect of gamification use. The same results were found in Putz et al. [40], a study conducted with 617 students using workshop design with gamification elements compared with non-gamified workshop designs, and in Sanchez et al. [45], a study that analyzed the used traditional quizzes and gamified online quizzes with 473 participants. Both reported the novelty effect.

1.3 | Learning outcomes

Despite several authors' high acceptance of gamification in educational contexts, studies should be analyzed integratively. There are many dispersed studies with several differences in their results. Although gamification impacts motivation, interest, and engagement, its effect on students' learning is not as straightforward.

In this scope, Sailer and Homner [42] performed a meta-analysis on the effects of gamification on cognition, motivation, and behavior. Their results showed significant minor effects of gamification in all three domains. However, some questions were not answered yet, mainly which factors contribute to the success of gamification since the results show high levels of heterogeneity, which is also in line with Marín et al.'s [28] study. The authors emphasize that studying whether the type of gamified elements used influences student performance is necessary. The authors studied students' learning performance levels when first-year engineering students used a gamified compiler compared or a non-gamified compiler to learn the C programming language. The results reveal that students obtained significantly better grades when the gamified platform, but further analysis is needed on the results and future data, especially on factors that were vital for success [28]. Murillo-Zamorano et al. [34] conducted a study where they verified the influence of gamification on student's knowledge. However, the research instruments were based on self-perception. So, it still unclear whether students indeed increase their knowledge through gamification.

Maryono et al. [30] conducted a systematic review of the effects of gamification in programming learning but with no emphasis on learning. The authors verified that using points and leaderboards can lead low-ranking students to stress and have a lower impact on highly complex programming content. Similar to Sailer and Homner [42] and Freitas and Silva's [22] studies, Maryono et al. [30] found an improvement in students' interest, motivation, and engagement in programming learning. The need for further research for clearly defined components that describe precise mechanisms by which gamification can affect specific learning processes and outcomes remains partially unresolved.

This study will help answer this question in the programming learning field, a discipline in continuous expansion from primary education (e.g., [25]). Through a multi-level meta-analysis it will be analyzed 1) how gamification affects programming learning and, 2) the impact of the most used game concepts, namely levels, points, badges, avatars, and leaderboards [47, 59] on knowledge acquisition.

2 | METHOD

The effects of using gamified mechanisms in programming learning, specifically in the students' performance, were analyzed first through the systematic review process lens and then by conducting a meta-analysis.

2.1 | Literature search

This study followed the PRISMA statement in the literature search phase, a strategy for reporting systematic reviews and meta-analyses [33]. The search was carried out on June 30, 2020, in the databases ERIC, Google Scholar, ISI Web of Knowledge, and Scopus since it was intended to cover a wide range of different outcomes, and in January 2022 in the database Scopus to update the previous data.

For ERIC, ISI, and Scopus databases, it was used the search terms "gamification" and ("programming" and ("learning" or "teaching" or "course")) or "learn to program" in the topic, with peer review and full text available. Since Google Scholar does not allow searching for a topic, it was used the search terms in the title. We used search terms concrete to ensure that they would return a set of results accurately. The search returned a total of 241 studies.

In January 2022, the results searches were updated to a new search in the Scopus database. The used search terms are as follows:

TITLE-ABS-KEY("gamification" and (("programming" and ("learning" or "teaching" or "course")) or "learn to program") AND (LIMIT-TO (PUBYEAR,2022) OR LIMIT-TO (PUBYEAR,2021) OR LIMIT-TO (PUBYEAR,2020))).

Eight authors of inaccessible articles in the databases were contacted through email. Four researchers have responded to the request.

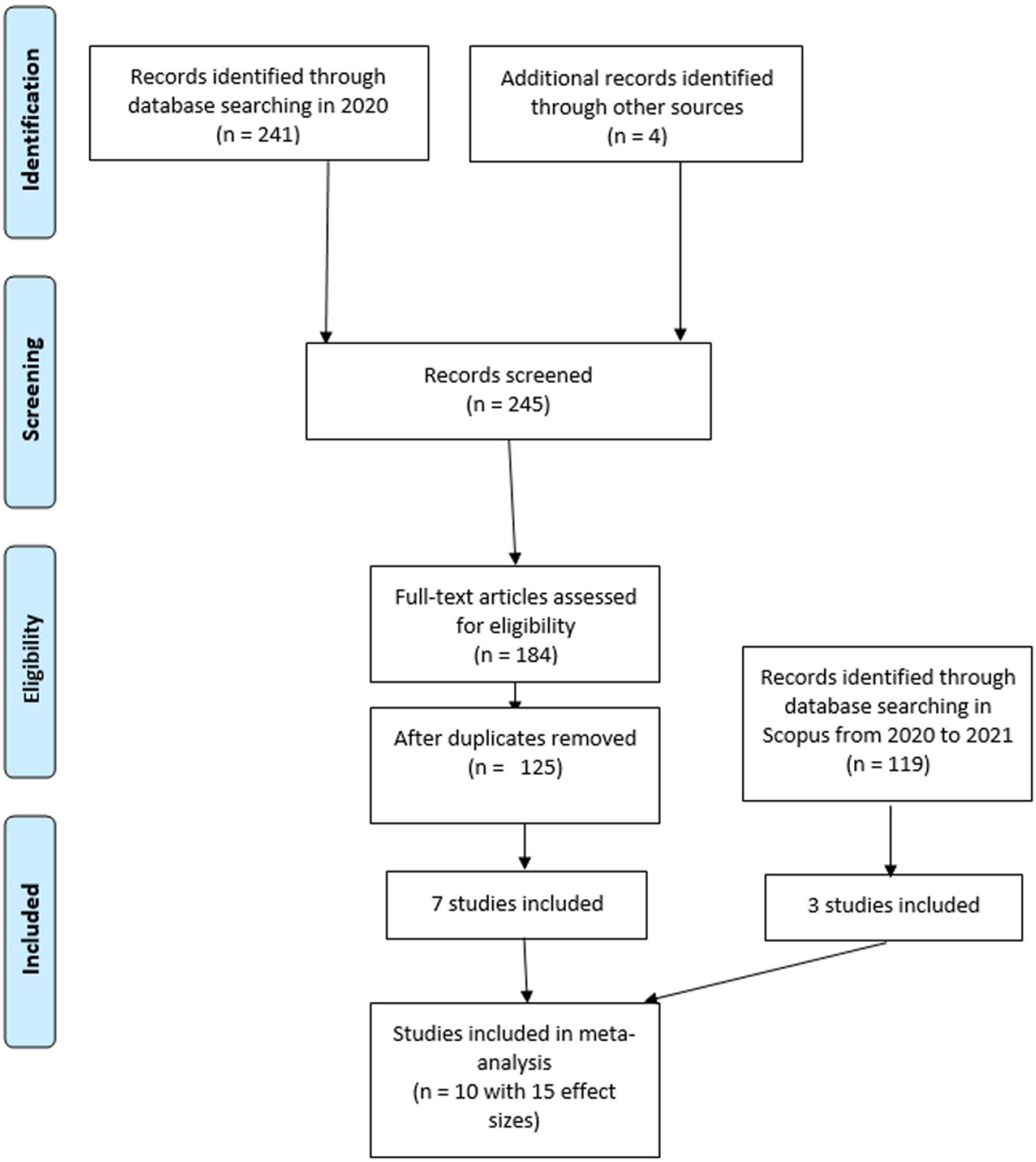

FIGURE 1 PRISMA flow diagram through gamification in programming learning.

2.2 | Inclusion and exclusion criteria

After gathering the articles provided from the search terms and removing the duplications ($n = 125$), the articles that met the following criteria were included:

1) The study uses gamification strategies.
2) It presents the effect size or quantitative data that allow its calculation.
3) The study was implemented in a K-12 sample or higher degree.
4) The performance was evaluated.
5) The study scope includes programming teaching or learning.
6) The study followed a quasi-experimental or experimental design.

Ten studies were selected after applying inclusion criteria, 10 studies were selected (seven from the first search and three from the second search in January 2022, which included studies between July 2020 and January 2022). Some of the 10 studies, presented more than one experiment, totaling 15 effect sizes (Figure 1).

2.3 | Moderators

The potential moderators were selected from the design of the studies and the gamification elements found in the studies. The impact of the sample size and the study's duration were the design's analyzed characteristics. The analysis of the gamification elements included the list of principal game elements presented in several cited studies: badges, points, leaderboards, levels, and avatars.

The analysis of each separated element and their combination allowed verifying which combinations are most effective in the scope of programming learning or whether the number of gamified elements is the most valuable.

Table 1 represents the definition of each moderator used in this study.

2.4 | Statistical analysis

After gathering the studies, the following steps consisted of calculating the effect size, comparing the studies, and proceeding to the meta-analysis.

In the first step, the studies were converted into a single measure, Cohen's d, using Lenhard and Lenhard's (2016) application. Cohen's d is a measure that represents an estimate of the number of subjects in the experimental group who are expected to exceed the medium value of the control group [8]. It has been used in meta-analysis in several areas, such as education [10], and medical research [18], among others. Cohen's d takes advantage of using an unbiased estimative of true population variability, regardless of the rejection of the null hypotheses, and it is also less biased than Hedges' g measure [8].

The second step was conducting the three-level meta-analysis, also known as multi-level meta-analysis [19].

TABLE 1 Moderators' description.

Moderator	Description
Badges	Although badges can be implemented in several ways [22], they are often a visual reward scoring system that displays the users' progress, leading users to perform specific activities to receive them [35].
Points	System reward by number with immediate effect when the user achieves some objective or correctly answers to a question [1]. Despite no studies researching the single use of points, Freitas and Silva [22], Murillo-Zamorano et al. [34] indicate that it is one of the most used game mechanisms.
Leaderboards	It presents the results ordered by achieves and/or points, ranking participants by their success in each activity. Freitas and Silva [22] stress that no conclusions can be made by using only leaderboards once most studies used at least two game mechanisms, making it difficult to understand the exact impact of each mechanism.
Levels	Release content to be unlocked progressively according to reaching objectives or points [35], increasing the difficulty and allowing the students to progress by acquiring distinct levels of knowledge [34, 17].
Avatars	Customized profile character. Krause et al. [26] also used extra accessories for avatars that students could collect and increase their customization.
Total	Sum of the used gamified elements, namely, badges, points, leaderboards, levels, and avatars.
Duration	Duration of the experience: 1 (less than 2 weeks); 2 (between 2 and 4 weeks); or 3 (more than 4 weeks)
Sample	Size of the sample: the sum of the participants in the control and experimental groups.

This approach was considered the most suitable once the treatment effect is estimated for each subgroup and outcomes, and it enables to test distinct moderator effects of study characteristics [58].

3 | RESULTS

Table 2 represents the characteristics of the selected studies, namely the calculated effect size, the number of participants (Sample), the duration of the study (Duration), the used design (Design), the game elements used in the experiments (Badges, Points, Avatar, Leaderboard, and Levels), and the total number of game elements.

3.1 | Overall model results

Table 3 presents the results of the pooled effect estimate. An estimate of 0.4308 represents a correlation of $r = .39$, which can be considered a medium correlation, showing that there is indeed an association between gamification and programming learning.

3.2 | Variance across the three levels

Table 4 shows the percentage of total variance by level. The value of I^2_{Level2}, representing the variance within studies, is around 19.6%. The value of I^2_{Level3}, which represents the variance between studies, is around 25.3%,

TABLE 2 Characteristics of the selected studies.

Study	Effect size (d)	Sample (n)	Duration	Design	Badges	Points	Avatar	Leaderboard	Levels	Total
Dominguez et al. [14] 2	0.021	45	3	Quasi-Experimental	No	Yes	No	Yes	No	2
Dominguez et al. [14] 1	0.215	46	3	Quasi-Experimental	No	Yes	No	Yes	No	2
Marin et al. [28] 2	0.07	410	3	Quasi-Experimental	No	Yes	No	Yes	No	2
Papadakis et al. [37]	0.147	30	1	Quasi-Experimental	No	Yes	Yes	No	No	2
Marin et al. [28] 1	0.383	410	3	Quasi-Experimental	Yes	Yes	No	Yes	No	3
Cadavid et al. [7] 1	0.49	45	3	Quasi-Experimental	Yes	Yes	No	Yes	Yes	4
Cadavid et al. [7] 2	0.51	45	3	Quasi-Experimental	Yes	Yes	No	Yes	Yes	4
Cadavid et al. [7] 3	0.51	45	3	Quasi-Experimental	Yes	Yes	No	Yes	Yes	4
Pontes et al. [37] 2	0.565	49	1	Quasi-Experimental	No	No	No	Yes	No	1
Pontes et al. [38]	0.661	60	1	Quasi-Experimental	Yes	Yes	No	Yes	Yes	4
Pontes et al. [37] 1	0.72	35	1	Quasi-Experimental	No	No	No	Yes	No	1
Shorn [48]	0.765	191	3	Quasi-Experimental	No	Yes	No	Yes	Yes	3
Beltran Morales et al. [4]	0.72	25	3	Quasi-Experimental	Yes	Yes	Yes	Yes	Yes	5
Tasadduq et al. [53]	0.18	46	3	Quasi-Experimental	Yes	Yes	No	No	Yes	3
Barriales et al. [5]	0.553	55	2	Quasi-Experimental	No	Yes	No	Yes	No	2

TABLE 3 Model results.

Estimate	0.4308
Standard error	0.0879
p-value	4.9021
t	0.0002
Confidence Interval (lower bound)	0.2423
Confidence Interval (upper bound)	0.6192

TABLE 4 Variance by level.

Level	% of total variance
1	55.07
2	19.63
3	25.29
Total I^2: 44.93%	

not much higher than within studies. Differences between studies can explain about a quarter of the total variance.

4 | DISCUSSION

4.1 | Choosing the best model

Choosing the model that represents a better fit for the data leads to the first analysis on the multi-level meta-analysis approach. A three-level model is helpful if the variability between the data is more noticeable than when using a two-level model. If the explanation is similar, Occam's razor principle is applied. So, the priority will be given to a Level 2 model because it is less complex and explains the data with similar rigor [19].

An ANOVA test was performed between the two models using the metafor package available for R language to check whether the nesting individual effect size studies would improve the model. Table 5 shows the difference between the Full model (the three-level, considering the dependence of the studies in the same article) and the Reduced model (the two-level model). As it can see, the values of Akaike (AIC) and Bayesian Information Criterion (BIC) are higher in the Full Model, which indicates that the reduced model allows for a better fit.

Despite the values of AIC and BIC being lower in the reduced model and therefore presenting a better performance, the results are very close between both. The likelihood ratio test is also not significant (p-value > .05), which empowers this analysis.

Besides, some of the selected studies present more than one experiment, indicating the presence of multiple effect sizes. So, these experiments cannot be independent [19]. For these reasons and because the three-level model represents the results better, the analysis will proceed with the three-level meta-analysis model.

4.2 | Effect sizes

Table 6 presents the effect size of the selected studies (d). The presence of a number in the study name means that there is more than one described experiment, and this identifier separates them. 95%-CI represents the confidence interval, and %W column shows the weight of each study in the meta-analysis under the random model.

From the random effect model, it can be observed that gamification indeed influences programming learning (Cohens' $d = 0.41$). Figure 2 and Table 6 show that more than half of the studies presented an effect size higher than 0.5, leading to the necessity to analyze the effect of moderators and verify whether there is a significant influence on the results that justify the heterogeneity in the results. No outliers were found.

4.3 | Moderators

A t-test on the β-weight of each moderator was performed to check whether the moderators impacted the differences between effect sizes [19]. The robustness of the meta-regression was verified by checking the intercorrelations between the moderators to verify if there were high correlations (Table 7).

There are high correlations between badges and levels, total and badges, and total and points (Figure 3). These results are in line with Mekler et al. [31] and Deterding et al.'s [12] studies when they referred that the combination of points, badges. and leaderboards is the most chosen group to gamified tasks or strategies.

As the total moderator shows a strong correlation between two variables, it was excluded from the subsequent analysis. However, badges and levels will remain in the analysis since they are among the most used gamified elements.

Regarding the analysis of each moderator, it was verified whether the use of avatars, leaderboards, levels, badges, and points would influence the magnitude of the effect of the 10 studies. Meta-regression revealed that only levels significantly impact the programming

TABLE 5 ANOVA test between two-level and three-level models.

Model	*Df*	AIC	BIC	AICc	logLik	LRT	*p*-val	QE
Full	3	9.58	11.49	11.98	−1.79			23.03
Reduced	2	8.03	9.31	9.12	−2.01	0.45	.5	23.03

Abbreviations: AIC, Akaike information criterion; ANOVA, analysis of variance; BIC, Bayesian Information Criterion.

TABLE 6 Effect sizes of the selected studies.

Study	*d*	95%-CI	%W (random)
Dominguez et al. [14] 2	0.0210	[−0.5474; 0.5894]	5.3
Dominguez et al. [14] 1	0.2150	[−0.3534; 0.7834]	5.3
Marin et al. [28] 2	0.0700	[−0.1260; 0.2660]	14.9
Papadakis et al. [37]	0.1470	[−0.5586; 0.8526]	3.8
Marin et al. [28] 1	0.3830	[0.1870; 0.5790]	14.9
Cadavid et al. [7] 1	0.4900	[−0.0980; 1.0780]	5.0
Cadavid et al. [7] 2	0.5100	[−0.0780; 1.0980]	5.0
Cadavid et al. [7] 3	0.5100	[−0.0780; 1.0980]	5.0
Pontes et al. [37] 2	0.5650	[−0.0034; 1.1334]	5.3
Pontes et al. [37] 1	0.7200	[0.0536; 1.3864]	4.1
Pontes et al. [38]	0.6610	[0.1514; 1.1706]	6.1
Shorn [48]	0.7650	[0.4710; 1.0590]	11.4
Beltran Morales et al. [4]	0.7200	[−0.0640; 1.5040]	3.2
Tasadduq et al. [53]	0.1800	[−0.3884; 0.7484]	5.3
Barriales et al. [5]	0.5530	[0.0042; 1.1018]	5.5

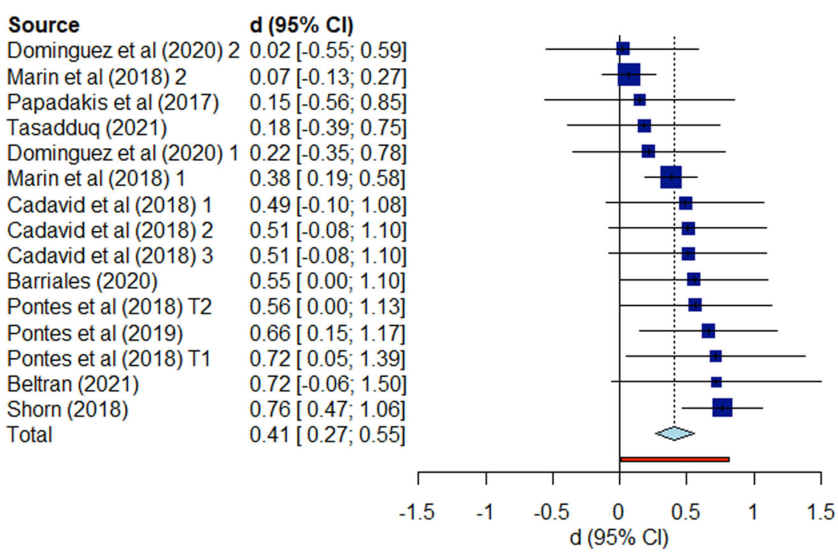

FIGURE 2 Forest plot with the effect size of the selected studies. CI, confidence interval.

TABLE 7 intercorrelation matrix for moderators.

	Badges	Points	Avatar	Leaderboard	Levels	Total
Badges	1.00000000	0.3668997	0.02620712	−0.02620712	0.73214286	0.8479461
Points	0.36689969	1.0000000	0.15384615	−0.15384615	0.36689969	0.6054055
Avatar	0.02620712	0.1538462	1.00000000	−0.42307692	0.02620712	0.2354355
Leaderboard	−0.02620712	−0.1538462	−0.42307692	1.00000000	−0.02620712	0.1009009
Levels	0.73214286	0.3668997	0.02620712	−0.02620712	1.00000000	0.8479461
Total	0.84794611	0.6054055	0.23543548	0.10090092	0.84794611	1.0000000

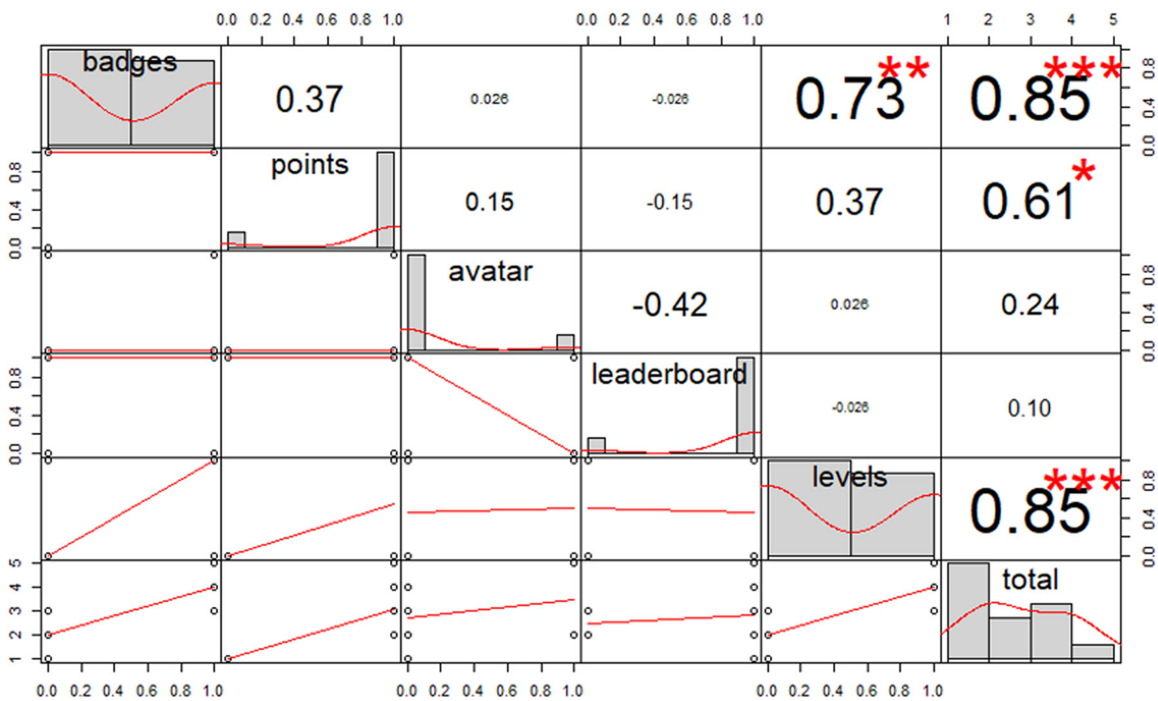

FIGURE 3 Intercorrelation between moderators.

learning ($p < .05$), explaining 21% of the heterogeneity. The use of badges, points, avatars, and leaderboards does not significantly influence the effect size ($p > .05$).

A permutation test was performed to assess the model fitting for levels moderator, to confirm whether there is a pattern in the analysis [19]. The permutation test confirmed that using levels influences programming learning (p-value $< .05$).

5 | CONCLUSIONS

Gamification has been integrated into the classroom to increase students' motivation, engagement, and time spent in gamified activities or to improve the learning process. The scattering of results, not consistently favorable, in different academic areas, and the number of gamified elements that can be used, led to the need to understand their effects on learning in greater depth. This article has focused on contributing to the discussion of this theme by analyzing the impact of gamification on programming learning, emphasizing the most used gamified elements.

The results of this meta-analysis revealed some details that have not yet been deeply researched and should be further addressed: although gamification positively affects programming learning, it seems to be exclusively related to using levels. This fact suggests some proximity to cognitive approaches that use difficulty levels as a nuclear basis, such as the 4 C/ID model, which has significantly impacted learning [9]. Besides, the need to reach a level only having completed the previous one, reinforces the need to reach the objectives of the current level. However, more empirical research is needed to verify this relation accurately.

The use of badges, points, avatars, and leaderboards showed no significant effect on programming learning, which suggests that these elements may be more beneficial to motivating or engaging the students, as seen in several studies (e.g., [59, 42]) than for learning purposes. Our results aline with Thomas and Baral's [54] study. The authors researched the effect of the flow experience, i.e. the "mechanism that creates the response from the stimulus in a gamified learning context" (p. 4), with three distinct instructional design through a within-subject experiment: traditional, game-based, and a gamified environment. The authors found that behavioral and emotional engagement increased during the gamified sessions but no significant change in cognitive engagement. Similar results can be found in Dichev and Dicheva's [13] work, related to a nonsignificant academic performance from using gamification.

5.1 | Limitations

This study has some limitations that should be addressed. Firstly, the concept of gamification in technological learning environments is broader than the areas covered here. Its effects on physical, emotional, social, and cognitive well-being, as well as its relation to mitigating potential health risks are areas of research that are emerging (e.g., [32]) and were not covered in this study.

The other main limitation is the generation of conclusions. Once the target field was programming learning, these conclusions may not be applied in other fields, where other skills and knowledge are required. So, the effects of using gamification can be distinct from other fields.

Finally, although many articles use gamification in programming learning, only 10 articles clearly showed clear evidence of its use in this field and presented precise data for meta-analysis. This gap has already been mentioned in Sailer et al. [43], indicating that empirical research on success, performance indicators, and progress had not yet been conducted. A three-level meta-analysis, also known as a multi-level meta-analysis, guaranteed that the studies that contributed with more than one effect size did not bias the results. This model controls this situation, assuming that the effect sizes are included in larger clusters [19] and increasing the number of contributions.

5.2 | Implications and future research

Among the game concepts studied, only levels positively influence programming learning. Although only the most used concepts were selected, this study puts the use of gamification in programming learning into perspective. Its use as an efficient strategy to enhance learning, which is not only related to motivation or engagement, needs a better understanding of the gamification process and how its effects can be enhanced to obtain the expected results. This study suggests that the use of levels can be one of the indicators of success. Thomas and Baral [54], in turn, explore the flow in the gamification process and suggest that this can help clarify the effect of gamification.

In future work, the use of more robust analyzes that can be integrated into meta-analyses is suggested to improve the robustness of this process. It is also essential to verify the effects of gamification in samples with younger participants. Since this meta-analysis focused mainly on youth and adults, it is necessary to understand whether it will have more effect on children.

Lastly, it is proposed to reinforce the use of levels in learning strategies that use gamification elements. The results of this meta-analysis suggest that similar to teaching strategies that use scaffolding [20], its application will improve programming learning.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Joana M. Costa https://orcid.org/0000-0001-7400-0132

REFERENCES

1. Y. Attali and M. Arieli-Attali, *Gamification in assessment: do points affect test performance*? Comput. Educ. **83** (2015), 57–63.
2. S. Azmi, N. A. Iahad, and N. Ahmad. 2016), Attracting students' engagement in programming courses with gamification. In 2016 IEEE Conference on e-Learning, e-Management and e-Services (IC3e, October). (112–115). IEEE.
3. S. Bai, K.F. Hew, M. Sailer, and C. Jia, *From top to bottom: how positions on different types of leaderboard may affect fully online student learning performance, intrinsic motivation, and course engagement*, Comput. Educ. **173** (2021). https://doi.org/10.1016/j.compedu.2021.104297
4. J. Beltran Morales, H. Sánchez, and M. Rico, *Aprendizaje divertido de programación con gamificación*, Revista Ibérica de Sistemas e Tecnologias de Informação (2021), no. 41, 17–33.
5. A. F. Barriales, J. V. Paragulla, and L. Andrade-Arenas, *Gamification as part of teaching and its influence on learning computational algorithms*. 2020 IEEE World Conference on

Engineering Education (EDUNINE), Bogota, Colombia, 2020, pp. 1–4. https://doi.org/10.1109/edunine48860.2020.9149510
6. A. Bogdanovych and T. Trescak, April *Teaching programming fundamentals to modern university students*, CSEDU (2016), no. 2, 308–317.
7. J. M. Cadavid and A. P. Corcho, *Competitive programming and gamification as strategy to engage students in computer science courses*, Revista ESPACIOS **39** (2018), no. 35.
8. J. E. Conboy, *Algumas medidas típicas univariadas da magnitude do efeito*, Análise Psicológ. **21** (2012), no. 2, 145–158.
9. J. M. Costa, G. L. Miranda, and M. Melo, *Four-component instructional design (4C/ID) model: a meta-analysis on use and effect*, Learn. Environ. Res. **25** (2022), no. 2, 445–463. https://doi.org/10.1007/s10984-021-09373-y
10. J. M. Costa, S. Moro, G. Miranda, and T. Arnold, *Empowered learning through microworlds and teaching methods: a text mining and meta-analysis-based systematic review*, Res. Learn. Technol. **28** (2020). https://doi.org/10.25304/rlt.v28.2396
11. K. Dcuervo-Cely, F. Restrepo-Calle, and J. J ramírez-Echeverry, *Effect of gamification on the motivation of computer programming students*, J. Inform. Technol. Educ.: Res. **21** (2022), 001–023. https://doi.org/10.28945/4917
12. S. Deterding, D. Dixon, R. Khaled, and L. Nacke, September *From game design elements to gamefulness: defining "gamification"*, Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments (2011), 9–15.
13. C. Dichev and D. Dicheva, *Gamifying education: what is known, what is believed and what remains uncertain: a critical review*, Int. J. Educ. Technol. Higher Educ. **14** (2017), no. 1, 9.
14. A. Domínguez, L. de-Marcos, and J.-J. Martínez-Herráiz, *Effects of competitive and cooperative classroom response systems on quiz performance and programming skills in a video game programming course*. Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education, 2020. https://doi.org/10.1145/3341525.3387393
15. L. Facey-Shaw, M. Specht, and J. Bartley-Bryan, October *Digital badges for motivating introductory programmers: qualitative findings from focus groups*, 2018 IEEE Front. Educ. Conf. (FIE), IEEE, 2018, pp. 1–7.
16. M. Featherstone and J. Habgood, *UniCraft: exploring the impact of asynchronous multiplayer game elements in gamification*, Int. J. Hum. Comput. Stud. **127** (2019), 150–168.
17. J. Fernandez-Rio, E. Heras, T. González, V. Trillo, and J. Palomares, *Gamification and physical education: viability and preliminary views from students and teachers*, Phys. Educ. Sport Pedagog **25** (2020), no. 5, 509–524. https://doi.org/10.1080/17408989.2020.1743253
18. A. B. Haidich, *Meta-analysis in medical research*, Hippokratia **14** (2010), no. Suppl 1, 29–37.
19. M. Harrer, P. Cuijpers, T.A Furukawa, and D. D. Ebert (2019). Doing meta-analysis in R: A hands-on guide. https://bookdown.org/MathiasHarrer/Doing_Meta_Analysis_in_R/. Accessed 04 January 2022; https://doi.org/10.1111/ejed.12363
20. J. Hattie, 250+ influences on student achievement, 2017. https://visible-learning.org/wp-content/uploads/2018/03/VLPLUS-252-Influences-Hattie-ranking-DEC-2017.pdf
21. A. Hoskey and P. S. M. Maurino, *Beyond introductory programming: success factors for advanced programming*, Inform. Syst. Educ. J. **9** (2011), no. 5, 61.
22. M. Jarnac de Freitas and M. Mira da Silva, *Systematic literature review about gamification in MOOCs*, Open Learn.: J. Open, Distance e-Learn. **38** (2023), no. 1, 73–95. https://doi.org/10.1080/02680513.2020.1798221
23. Y. Jia, B. Xu, Y. Karanam, and S. Voida (2016). *Personality-targeted gamification: a survey study on personality traits and motivational affordances*. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (2001–2013).
24. A. Kesler, T. Shamir-Inbal, and I. Blau, *Active learning by visual programming: pedagogical perspectives of instructivist and constructivist code teachers and their implications on actual teaching strategies and students' programming artifacts*, J. Educ. Comput. Res. **60** (2022), no. 1, 28–55.
25. S. C. Kong and Y. Q. Wang, *Investigating primary school principals' programming perception and support from the perspective of reasoned action: a mixed methods approach*, Comput. Educ. **172** (2021), 104267. https://doi.org/10.1016/j.compedu.2021.104267
26. M. Krause, M. Mogalle, H. Pohl, and J. J. Williams (2015). *A playful game changer: Fostering student retention in online education with social gamification*. In Proceedings of the Second (2015) ACM conference on Learning@ Scale (95–102). https://doi.org/10.1145/2724660.2724665
27. T. Lehtonen, T. Aho, E. Isohanni, and T. Mikkonen (2015). (November) *On the role of gamification and localization in an open online learning environment: javala experiences*. In Proceedings of the 15th Koli Calling Conference On Computing Education Research (50–59).
28. B. Marín, J. Frez, J. Cruz-Lemus, and M. Genero, *An empirical investigation on the benefits of gamification in programming courses*, ACM Trans. Comput. Educ. **1** (2018), no. 19. https://doi.org/10.1145/3231709
29. J. M. Costa and G. L. Miranda, *Using alice software with 4C-ID model: effects in programming knowledge and logical reasoning*, Inform. Educ. **18** (2019), no. 1, 1–15. https://doi.org/10.15388/infedu.2019.01
30. D. Maryono, S. Budiyono, M. Sajidan, and M. Akhyar, *Implementation of gamification in programming learning: literature review*, Int. J. Inform. Educ. Technol. **12** (2022), 1448–1457.
31. E. D. Mekler, F. Brühlmann, A. N. Tuch, and K. Opwis, *Towards understanding the effects of individual gamification elements on intrinsic motivation and performance*, Comput. Human. Behav. **71** (2017), 525–534.
32. C. Melo, L. Madariaga, M. Nussbaum, R. Heller, S. Bennett, C. C. Tsai, and J. van Braak, *Educational technology and addictions*, Comput. Educ. **145** (2020), 103730. https://doi.org/10.1016/j.compedu.2019.103730
33. D. Moher, A. Liberati, J. Tetzlaff, and D. G. Altman, *Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement*, PLoS Med. **6** (2009), no. 7, e1000097.
34. L. R. Murillo-Zamorano, J. Á. López-Sánchez, M. J. López-Rey, and C. Bueno-Muñoz, *Gamification in higher education: the ECOn+ star battles*, Comput. Educ. **194** (2023), 104699. https://doi.org/10.1016/j.compedu.2022.104699

35. R. P. Oliveira, C.G. Souza, A.C. Reis, and W.M. Souza, *Gamification in e-learning and sustainability: a theoretical framework*, Sustainability **13** (2021), no. 21, 11945.
36. E. T. OPRIȘ, É. Bálint-Svella, and I. Zsoldos-Marchiș, *Prospective preschool and primary school teachers' knowledge and opinion about gamification*, Acta Didactica Napocensia **14** (2021), no. 1, 104–114.
37. S. Papadakis and M. Kalogiannakis, *Using gamification for supporting an introductory programming course. The case of classcraft in a secondary education classroom*. In Interactivity, Game Creation, Design, Learning, and Innovation: 6th International Conference, ArtsIT 2017, and Second International Conference, DLI 2017, Heraklion, Crete, Greece, October 30–31, 2017, Proceedings 6, Springer International Publishing, 2018, pp. 366–375.
38. R. G. Pontes, D. D. Guerrero, and J. C. Figueiredo, *Analyzing gamification impact on a mastery learning introductory programming course*. In Proceedings of the 50th ACM technical symposium on computer science education, 2019, pp. 400–406.
39. R. G. Pontes, K. H. M. Medeiros, D. D. S. Guerrero, and J. C. A. de Figueiredo, *Analyzing the impact of leaderboards in introductory programming courses' short-length activities*. 2018 IEEE Frontiers in Education Conference (FIE), San Jose, CA, USA, 2018, pp. 1–9. https://doi.org/10.1109/fie.2018.8658937
40. L.M. Putz, F. Hofbauer, and H. Treiblmaier, *Can gamification help to improve education? Findings from a longitudinal study*, Comput. Human. Behav. **110** (2020), 106392. https://doi.org/10.1016/j.chb.2020.106392
41. L. Rodrigues, F. D. Pereira, A. M. Toda, P. T. Palomino, and M. Pessoa, *Gamification suffers from the novelty effect but benefits from the familiarization effect: findings from a longitudinal study*, Int. J. Educ. Technol. Higher Educ. **19** (2020), no. 13. https://doi.org/10.1186/s41239-021-00314-6
42. M. Sailer and L. Homner, *The gamification of learning: a meta-analysis*, Educ. Psychol. Rev. **32** (2020), 77–112. https://doi.org/10.1007/s10648-019-09498-w
43. M. Sailer, J. Hense, J. Mandl, and M. Klevers, *Psychological perspectives on motivation through gamification*, Interac. Design Architect. J. **19** (2014), 28–37.
44. M. Sailer, J. Hense, H. Mandl, and M. Klevers, *Fostering development of work competencies and motivation via gamification*, Competence-based vocational and professional education, Springer, Cham, 2017, pp. 795–818.
45. D. R. Sanchez, M. Langer, and R. Kaur, *Gamification in the classroom: examining the impact of gamified quizzes on student learning*, Comput. Educ. **144** (2020), 103666. https://doi.org/10.1016/j.compedu.2019.103666
46. S. Schöbel, M. Saqr, and A. Janson, *Two decades of game concepts in digital learning environments – a bibliometric study and research agenda*, Comput. Educ. (2021), 173. https://doi.org/10.1016/j.compedu.2021.104296
47. A. Shafie and Z. Abdullah (2020. (March) *Gamification in learning programming language*. In SA Conference Series: Industrial Revolution 4.0 (Vol. **1**, No. 1, pp. 181–187).
48. S. P. Shorn, *Teaching computer programming using gamification*. In Proceedings of the 14th International CDIO Conference, Kanazawa Institute of Technology, Kanazawa, Japan, 2018.
49. L. Shürmann and C. Quaiser-Pohl, *Digital badges affect need satisfaction but not frustration in males in higher education*, Comput. Educ. **182** (2022). https://doi.org/10.1016/j.compedu.2022.104484
50. R. Smiderle, L. Marques, and S. Rigo (2019). *Studying the impact of gamification on learning and engagement of introverted and extroverted students*. In 2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT) (Vol. 2161, 71–75). IEEE.
51. R. Smiderle, S.J. Rigo, L.B. Marques, J.A. Peçanha de Miranda Coelho, and P.A. Jaques, *The impact of gamification on students' learning, engagement and behavior based on their personality traits*, Smart Learn. Environ. **7** (2020), no. 1, 3.
52. R. R. Sousa and F. T. Leite, *Usando gamificação no ensino de programação introdutória*, Brazilian J. Dev. **6** (2020), no. 6, 33338–33356.
53. M. Tasadduq, M.S. Khan, R.M.A. Nawab, M.H. Jamal, and M.T. Chaudhry, *Exploring the effects of gamification on students with rote learning background while learning computer programming*, Comput. Appl. Eng. Educ. **29** (2021), 1871–1891. https://doi.org/10.1002/cae.22428
54. N. J. Thomas and R. Baral, *Mechanism of gamification: role of flow in the behavioral and emotional pathways of engagement in management education*, Int. J. Manage. Educ **21** (2023), no. 1, 100718. https://doi.org/10.1016/j.ijme.2022.100718
55. M. Thongmak (2017). *Gamifying the first programming class: outcomes and antecedents of continued engagement intention*. In Proceedings of the International Conference on Information Systems Education and Research AIS SIGED.
56. F. Tomaselli, O. Sanchez, and S. Brown, *How to engage users through gamification: the prevalent effects of playing and mastering over competing*, In Thirty Sixth International Conference on Information Systems (2015), 1–16.
57. M. Urh, G. Vukovic, E. Jereb, and R. Pintar, *The model for introduction of gamification into e-learning in higher education*, Proc.-Soc. Behav. Sci. **197** (2015), 388–397.
58. W. Van den Noortgate, J.A. López-López, F. Marín-Martínez, and J. Sánchez-Meca, *Three-level meta-analysis of dependent effect sizes*, Behav. Res. Methods. **45** (2013), 576–594. https://doi.org/10.3758/s13428-012-0261-6
59. M. Venter, April) *Gamification in STEM programming courses: State of the art*, In. 2020 IEEE Global Engineering Education Conference (EDUCON). IEEE (2020), 859–866.

AUTHOR BIOGRAPHY

Joana M. Costa, PhD, is an assistant professor and integrated researcher at Iscte, Portugal. She holds the position of Director of the 1st cycle degree in Digital Educational Technologies at Iscte. Her research is primarily focused on the intersection of technology and education, with a specific emphasis on computer education, technology-enhanced teaching and learning, educational technology, and information systems in education. Joana has published 15 scientific articles published in indexed journals and has made active contributions to several projects in her field of study.

How to cite this article: J. M. Costa, *Using game concepts to improve programming learning: a multi-level meta-analysis*, Comput. Appl. Eng. Educ. 2023;31:1098–1110.
https://doi.org/10.1002/cae.22630