CBIE 2023
CONGRESSO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO
Uma escola para o futuro: Tecnologia e conectividade a serviço da educação

Motivation no Jutsu: Exploring the Power of Badge Aesthetics in Gamified Learning

Luiz Rodrigues[1], Armando M. Toda[2], Paula T. Palomino[3], Ana C. T. Klock[4], Anderson Avila-Santos[1]

[1]UniSENAI PR - Londrina - Brazil

[2]University of São Paulo - São Carlos - Brazil

[3]São Paulo State College of Technology - FATEC - Brazil

[4]Gamification Group - Tampere University - Tampere - Finland

luiz.rodrigues@sistemafiep.org.br

Abstract. *Gamification researchers have discussed the relevance of connecting game elements' visual appearance to students' feelings as a way to enhance learning outcomes. However, there is limited evidence from experimental studies on how gamification affects learning experiences. This study tackles this gap by investigating how students' feelings toward the visual appearance of badges affect their learning experiences in a gamified educational system. We conducted an experimental study (n = 19) comparing non-thematic badges to thematic ones featuring characters from a popular animation brand (i.e., Naruto), while controlling for the moderating effect of students' enthusiasm towards such brand. The results suggested that connecting gamification designs to students' feelings through visually appealing game elements can enhance their learning experiences, possibly because most students enjoy the anime and do not like non-thematic badges. Nevertheless, the results also raised concerns about the brand badges, such as distracting students. Thus, this paper informs practitioners and researchers on the value of designing gamification so that its aesthetics connect to students' feelings and raises concerns on potential issues to be considered and further explored.*

1. Introduction

Gamification design is a frequently discussed issue, which involves the game elements, along with their aesthetics and mechanics, applied in a non-gaming context [Deterding et al. 2011]. From a theoretical perspective, researchers have proposed several frameworks to support practitioners in better defining their gamification designs [Mora et al. 2017]. From a practical perspective, empirical studies often concern evaluating gamification designs to understand and justify their outcomes, especially when the results differ from the expected ones [Loughrey and Broin 2018, Toda et al. 2018].

Due to its relevance, the literature argues a possible explanation for gamification design issues is how they are commonly defined. As discussed in the literature review by [Koivisto and Hamari 2019], gamification designers often focus on selecting patterns of game elements to be available in the gamified environment. Accordingly, studies on gamification design are frequently centered on comparing the different game elements included in each design [Huang et al. 2020, Bai et al. 2020]. Furthermore, even in the case

of personalized gamification, which aims to offer different gamification designs depending on the user's characteristics and context, studies are mostly concerned with changing the game elements available, whereas few have explored personalizing game elements' aesthetics [Klock et al. 2020, Rodrigues et al. 2020, Hallifax et al. 2019].

In game design, the inspiration for gamification [Deterding et al. 2011], aesthetics is a core component of a concept named *juiciness*. Juiciness has emerged as a key element for enhancing player engagement and satisfaction, which incorporates audiovisual feedback that provides a sense of responsiveness, smoothness, and delight to the player's interactions within a game [Hicks et al. 2019a]. Particularly in the context of gamification applied to education, where the aesthetics of the learning experience play a vital role, the incorporation of juiciness can offer promising opportunities to enhance learner engagement and motivation [Hicks et al. 2019b]. However, while juiciness has been extensively explored and integrated into game design practices, its application within the field of gamification remains relatively unexplored, which corroborates the scarcity of research on the aesthetics of gamification's game elements.

Despite little research, the literature provides promising insights that changing game elements' aesthetics might benefit learning experiences by evoking nostalgia, as well as connecting to one's affective memory and personal preferences [Cardoso et al. 2017, Chou 2019]. Still, in gamification's context, empirical evidence on how students' feelings regarding game elements' aesthetics affect their learning experiences is limited to studying badges targeting the learning activity's subject [Pereira et al. 2023] and student's affective memories toward the Pokémon brand [Rodrigues et al. 2022a]. Nevertheless, previous research has not contributed empirical evidence from experimental studies on i) how badges with aesthetics appealing to students' feelings compare to those that do not and ii) how this appeal affects gamification outcomes. Therefore, there is a need to fulfill this lack to inform gamification design and support (or not) efforts to make game elements visually appealing to users.

Based on that, this paper's goal is to understand how badges from a specific brand affect learning experiences, compared to non-thematic badges, depending on students' feelings regarding that brand. For this, we conducted a mixed-methods, experimental study [Barbosa et al. 2021] wherein we implemented non-thematic badges through *medals* - as those are among the most used in recent literature [Huang et al. 2020, Koivisto and Hamari 2019] - and explored the Naruto brand, given its popularity with young people, to implement thematic badges. Thus, this paper contributes empirical evidence that informs practitioners and researchers on the potential of employing gamification designs that connect to students' feelings as a way to improve learning outcomes based on badges' aesthetics.

2. Related Work

Most often, studies aiming to understand how gamification designs affect learning experiences are centered on the game elements each design features [Huang et al. 2020, Bai et al. 2020]. On the one hand, that focus is supported by research demonstrating that different people prefer and are motivated by distinct game elements (e.g., [Palomino et al. 2023, Tondello et al. 2019]). On the other hand, the literature also shows that designing game elements so that they are visually appealing to users has the potential

to improve their experiences.

For instance, [Linehan and Kirman 2017] argues that Pokémon Go's success relates to its visual Pokémon layer due to players' feelings (e.g., nostalgia) regarding the brand. Accordingly, research shows that nostalgia is related to affective memory, which in turn has a prominent role in the interactions and experiences people have with a product [Cardoso et al. 2017]. Similarly, [Chou 2019] suggests gamification designs should be connected to users' past memories and personal interests.

Despite the potential of game elements' aesthetics for improving gamification outcomes, little research has explored this line. For instance, [Kao and Harrell 2018] investigated how different badge aesthetics affect people's experiences when playing and designing levels for an educational game. They particularly found that a specific aesthetic (i.e., role models, such as Einstein) was effective in improving intrinsic motivation, arguing the key reason for such a result is that the visual of the badge was relevant to users. Notably, the study concerns receiving badges within an educational game, whereas this paper explores them in gamification applied to education [Deterding et al. 2011].

Differently, [Pereira et al. 2021] analyzed the usage of thematized badges in the educational domain. The study proposes and exemplifies several badges targeting human-computer interaction education. The idea is to use badges with aesthetics based on distinguished researchers of the field aiming to connect them to the subjects being taught and, consequently, improve their relevance for students. This approach has been expanded and applied to other courses, such as Introductory Computing [Pereira et al. 2023]. Nevertheless, its application is limited to an experience report and lacks an experimental analysis of how it compares to badge visuals not targeting the course's subject.

Instead of badges related to the learning activity, [Rodrigues et al. 2022a] studied the role of badges aimed at students' affective memories. They conducted a usability study to understand how students' experiences differ depending on the badge aesthetic they received (Medals or Pokémons), assuming the participants would enjoy the Pokémon ones based on affective memories related to the brand (e.g., from the anime/games). They found students had better experiences with the Pokémon badges, and qualitative insights suggested the reason for that result is the brand being related to participants' affective memories as well as fostering intrinsic motivation through curiosity and relatedness. Importantly, the study highlights those explanations demanded future investigations to empirically support and test their generalization.

Based on that review, Table 1 summarizes related work's main characteristics. Although prior research has explored how badges' aesthetics affect user experience, previous studies have explored it outside the gamification context [Kao and Harrell 2018]; designed badges' visuals targeting the task's subject, not users [Pereira et al. 2023]; or demand empirical support [Rodrigues et al. 2022a]. Therefore, despite arguments on the value of badges' visual appearance to gamification's outcomes [Chou 2019, Linehan and Kirman 2017], there is a lack of empirical evidence on how badges connected to students' feelings compared to those that do not. Thus, this paper addresses that gap with an experimental study investigating how students' feelings regarding badges' aesthetics affect their experience with a gamified system.

Table 1. Summary of related work.

Reference	Context	Badge target	Study Method
[Kao and Harrell 2018]	Games	Users	Experimental
[Rodrigues et al. 2022a]	Gamification	Users	Usability Test
[Pereira et al. 2023]	Gamification	Task Subject	Experience Report
This study	Gamification	Users	Experimental

3. Method

To achieve our goal, we conducted an experimental, laboratory study using the mixed-methods approach [Barbosa et al. 2021]. The study is based on a single factor (i.e., gamification) with two levels (i.e., thematic - Naruto; non-thematic - Medals) and the within-subject design, as related research has recommended [Klock et al. 2020]. Accordingly, we used counterbalancing to mitigate order effects: participants were randomly attributed to either use the non-thematic version, then the thematic one, or vice-versa [Wohlin et al. 2012]. On the one hand, the non-thematic version aimed to replicate the standard gamification design as medals are one of the aesthetics used the most [Koivisto and Hamari 2019, Huang et al. 2020]. On the other hand, the thematic (Naruto) version aimed to be aligned with participants' interests. Based on that setting, this study tested the following hypotheses:

- **H1:** Student intrinsic motivation is higher for thematic badges compared to non-thematic ones.
- **H2:** The higher the student's feelings towards the Naruto brand, the more they prefer the thematic badges compared to the non-thematic ones.

Moreover, to further understand the results related to **H1** and **H2**, we also sought to answer the following research question (**RQ**): *What reasons explain how students' experiences with thematic badges compare to those with non-thematic ones?* In light of this study's overview, the remaining of this section describes the study participants, procedure, experimental task, gamification versions, instruments, and data analysis procedure.

Nineteen undergraduate students participated in this study. Those were selected by convenience sampling [Wohlin et al. 2012] as they were enrolled in the Data-oriented Statistics class (Software Engineering course from a Brazilian, private, southern institution) where the first author was lecturing during the 2022-2 term. Of those, 18 identified as men and one as woman. Their average age was 20 (Standard Deviation, SD = 3). To be a study participant, the procedure was as follows. First, the student had to accept the invitation and agree with the consent form. Second, the person had to complete the measure that captured an objective estimation of how they felt regarding Naruto. Third, they completed a quiz using each gamification version at a time (the order was randomized, as described above). After using each version, the participant completed the measure that captured their intrinsic motivation (see the instruments below). Finally, the student participated in a non-structured interview focused on their experience with the gamification versions.

The experimental task was inspired by the related work of [Rodrigues et al. 2022a]. The task consisted in completing a quiz based on five multiple-choice items, wherein each item featured four alternatives; a single one was

correct (see Figure 1a). The items' theme was Waste Recycling, which we chose due to the subject's relevance for sustainable development. To deploy this quiz, we used an interactive prototype developed on AdobeXD[1]. In the prototype, participants found information regarding the study goals, instructions on how to use it, the five multiple-choice items, and a thank you page. Importantly, this task was limited to five items to maintain the feasibility of prototyping it.

(a) Sample of a quiz item.

(b) Sample of the thematic (Naruto) version.

Figure 1. Screenshots of the quiz prototype[2] used in the experimental task.

In gamifying the quiz, our goal was to motivate the students to continue using it. Therefore, we used badges, an implementation of the Acknowledgment game element, as a way to provide positive feedback every time the student correctly answered a quiz item [Toda et al. 2019]. The rationale is that the feedback would motivate the person to continue completing items to receive more badges. Based on that, we designed two gamification versions, which only differ because the non-thematic one rewards users with medals and trophies, whereas the thematic one acknowledges users' correct answers with Naruto avatars (see Figure 1b). Importantly, both versions always show the badges at the screen's top and react to each user's answer to communicate the gamification's goal.

To understand participants' experiences with the gamification versions, we used a combination of quantitative and qualitative data. To test **H1**, we quantitatively measured participants' intrinsic motivation, as it relates to both behavior (e.g., continue using) and learning experiences (e.g., learning gains) [Ryan and Deci 2017]. For this, we used the interest/enjoyment scale of the Brazilian Portuguese version of the Intrinsic Motivation Inventory [Pedro 2016]. To enable testing **H2**, we used the Emotional Engagement Scale

[1]https://helpx.adobe.com/br/xd/get-started.html

[2]This prototype was created using free icons from https://icons8.com/icon/set/characters/color.

[So et al. 2014] to measure participants' enthusiasm regarding Naruto, similar to prior research [Xi and Hamari 2019]. Both instruments are based on a seven-point Likert scale and demonstrated to be reliable based on Cronbach's Alpha (0.91 for intrinsic motivation; 0.94 for enthusiasm). To answer our **RQ**, we captured participants' subjective experiences based on the non-structured interviews conducted after everyone used both versions [Barbosa et al. 2021]. Those aimed to foster discussion among participants, reveal their opinions regarding the gamification versions, and were driven by the following question: *What did you think of the two prototype versions you just used?*

The data analysis process was as follows. To test **H1**, given the within-subject design, we relied on participants' intrinsic motivation *difference* (i.e., subtracting answers for the non-thematic version from those of the thematic one) and used it as the dependent variable in a one-sample t-test. To test **H2**, we used a linear regression wherein the intrinsic motivation difference and enthusiasm were the dependent and independent variables, respectively. Because both tests have a number of assumptions, we followed literature recommendations and also tested **H1** and **H2** using robust alternatives [Wilcox 2017]. As recommended, we used R's packages *WRS2* and *MASS*, and run the *onesampb* and *rlm* functions as the robust alternatives for t-test and linear regression, respectively. In all analyses, we adopted the standard 95% confidence level. Concerning the qualitative data, the first author reviewed the group discussion, took notes on comments related to the gamification versions, and grouped them into the pros and cons of the thematic version to answer our RQ. For both groups, the researcher sought and identified recurrent themes based on similar comments, which are supported by relevant quotes to maximize their validity when presented and discussed in Section 4.

4. Results

First, we hypothesized the thematic version would motivate students more than the non-thematic one (**H1**). Figure 2 shows these results, demonstrating students' intrinsic motivation related to the thematic version (Mean, M = 5.15; Standard Deviation, SD = 1.33) was higher than that of the non-thematic version (M = 4.44; SD = 1.31). This difference (M = 0.714; SD = 1.12) is statistically significant according to both standard (p-value, p = 0.01; Estimate, E = 0.71; Confidence Interval, CI = [0.17 - 1.25]) and robust ($p < 0.01$; E = 0.61; CI = [0.19 - 1.11]) tests. According to Cohen's D, the difference was 0.63, which is considered a moderate to large effect size [Kotrlik and Williams 2003]. Thereby, these results support the hypothesis that students feel more intrinsically motivated by the thematic version, compared to the non-thematic one, indicating a moderate to large difference.

Second, we hypothesized there is a positive relationship between one's feelings towards Naruto and the extent to which they prefer the thematic version compared to the non-thematic one (**H2**). Overall, participants' enthusiasm for Naruto was moderated (M = 3.68, SD = 1.77), given the seven-point Likert scale. As shown in Figure 3, we found enthusiasm was a statistically significant predictor of how much students prefer the thematic version, compared to the non-thematic one, according to both standard (E = 0.31; Standard Error, SE = 0.13; t-statistic, t = 2.32; p = 0.03) and robust (E = 0.27, SE = 0.13, t = 2.16) regression analyses. Accordingly, the regression line was statistically significant (F-statistic (degrees of freedom), F(1, 17) = 5.39; p = 0.03;), yielding an adjusted R^2 of 0.20, which is considered a moderate to large effect size [Kotrlik and Williams 2003].

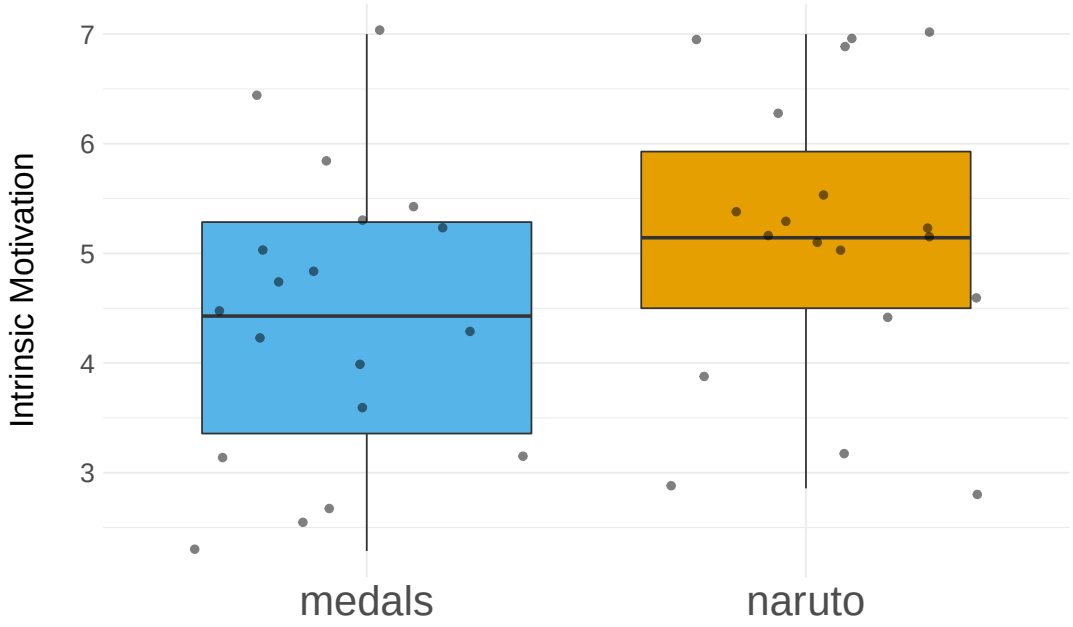

Figure 2. Students' intrinsic motivation from using our quiz prototype depending on whether they received medals (non-thematic) or Naruto (thematic) badges.

Therefore, these results confirm the hypothesis that the more one's enthusiasm for Naruto, the more they feel intrinsically motivated by its badges compared to the non-thematic one.

To further understand these findings, we analyzed students' comments to identify the pros and cons of the thematic version (**RQ**). Regarding reasons for feeling more motivated by the thematic version, students mentioned they enjoy the Naruto anime (e.g., *"I thought it was way more interesting [than the medal one]"*; *"It was interesting because you had the [Naruto] characters [...]"*; *"It was good because it's something we like"*) as well as have issues with the medal version (e.g., *"Because everyone's used to medals already, you know?"*, *"I thought it was kind of simple, didn't want to answer [the quiz]"*). Also, participants' reinforced the role of one's feelings toward the badges on how they experience them by arguing that *"if it [the badge theme] wasn't Naruto - it was One Piece - it would be more interesting"* and that *"it [the quiz] would be better with a Lightning McQueen version because I don't like Naruto very much"*. Hence, students' subjective experiences suggest that the pros of the thematic version compared to the non-thematic one are enjoying the Naruto anime, not liking medals, and liking the badges' visuals.

Importantly, we also found issues with the thematic version. For instance, some students argued they saw no difference between quiz versions (e.g., *"I identified the badges there, but [they were] the same thing to me"*; *"It was the same, the medals and the Naruto [badges]"*). Participants also raised concerns related to performance in the quiz. One of them mentioned that they missed more quiz items in the Naruto version due to the *"anxiety to get it right, you know?"* and justified that *"the [anime] badge incites you to get it [the item] right because you want to see it [the badge]"*. Another student claimed that they *"were paying attention to the quiz, not the badges"* in the medal version, but in the Naruto version they *"didn't care about the quiz anymore."* Additionally, a

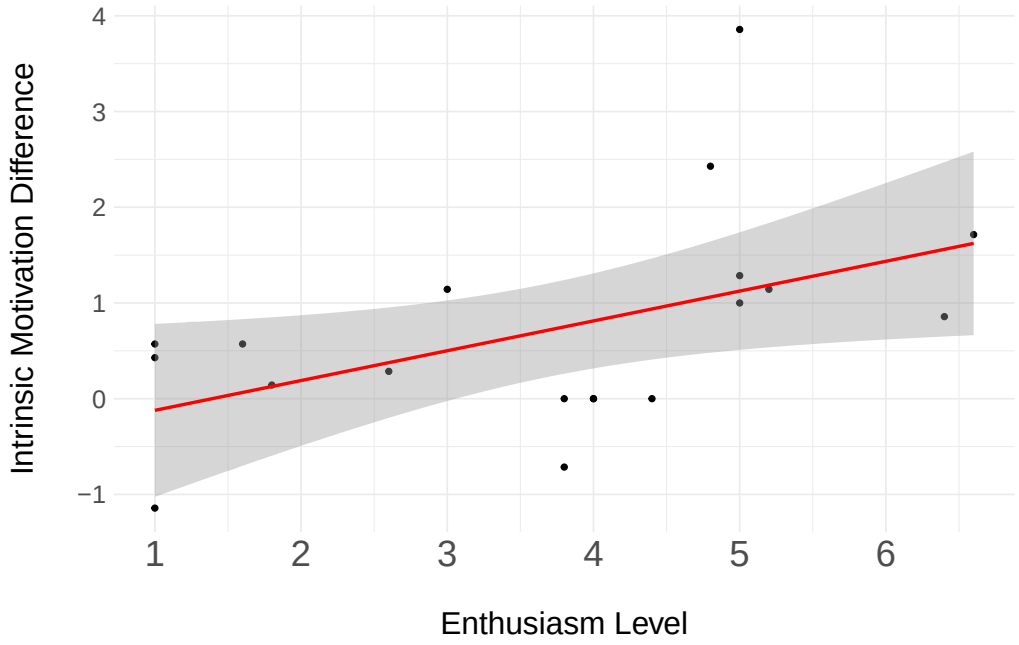

Figure 3. Relationship between students' enthusiasm with Naruto and the difference in the extent to which they feel intrinsically motivated to complete a quiz receiving Naruto badges compared to receiving medals.

student mentioned that they *"didn't like [the Naruto version] very much, [because] it gets too related to a single audience"* and another one argued that *"the person who doesn't watch [Naruto] wouldn't be very interested in finding out who is the character [of the badge]"*. Thereby, concerns related to the thematic version include being indifferent to or unfamiliar with the brand and distraction issues.

5. Discussion

In summary, our findings provide promising evidence for designing badges with visuals that appeal to users' feelings, despite also raising important considerations. On the one hand, we found quantitative and qualitative evidence supporting the value of employing gamification designs that connect to students' feelings based on badges' aesthetics. Quantitatively, we observed that learners who reported more enthusiasm for the Naruto brand reported significantly higher intrinsic motivation for the Naruto-based, thematic badges compared to the non-thematic badges (i.e., medals). Qualitatively, students expressed their enjoyment of the thematic badges due to their affinity for the anime, highlighting the positive impact of the brand connection. This finding corroborates research arguing that badges with visuals relevant to users can improve intrinsic motivation [Rodrigues et al. 2022a], as well as the idea of expanding juiciness benefits from games to gamification applied to education [Hicks et al. 2019b]. Therefore, this paper expands the literature with empirical evidence that confirms prior insights that designing gamification targeting users' feelings is valuable to improve its outcomes.

On the other hand, we also found concerns regarding the use of badges based on a specific brand. Particularly, some participants reported that they liked the thematic badges to the extent that they got distracted from the learning task, whereas others highlighted those who do not enjoy the brand might be indifferent to its badges. Past re-

search shows that gamification applied to education is especially challenging because it might lead to undesired behaviors that harm learning, such as gaming the system [Rodrigues et al. 2022b]. Similarly, studies have discussed that the same design is unlikely to work for everyone [Palomino et al. 2023, Tondello et al. 2019]. Thereby, while supporting the value of connecting game elements to users' feelings, we also expand the literature by demonstrating that this approach might also jeopardize learning and that exploring it based on a single brand does not fit all.

Based on that context, this paper contributes to the design of gamified learning by exploring aesthetics rather than different game elements. Most often, gamification studies have been concerned with comparing different (sets of) game elements to understand which of them might be associated with better outcomes [Bai et al. 2020, Huang et al. 2020]. In contrast, although the literature has provided interesting insights supporting connecting game elements to users' affective memories [Rodrigues et al. 2022a, Chou 2019, Linehan and Kirman 2017], there is a lack of understanding of how this approach compares to using game elements based on standard visuals. Thus, our contribution is empirical evidence from an experimental study that advances the understanding of how making game elements visually appealing to users' interests affects gamification outcomes.

Given our contribution, this study has two main implications. First, our results suggest that gamification designs that connect to students' feelings, such as utilizing badges with visuals from brands meaningful to users, have the potential to improve learning experiences. This finding emphasizes the importance of considering users' interests in gamification design from a visual perspective. Particularly, the findings suggest a positive effect on intrinsic motivation, which has a strong relationship to learning gains [Hanus and Fox 2015, Rodrigues et al. 2021]. Hence, our empirical evidence informs that this approach might ultimately help maximize learning. Second, our results suggest that using the same brand might not work for those not interested in it. This finding corroborates the literature in the sense that one size does not fit all, expanding it to the context of game elements' aesthetics. Thereby, we provide empirical evidence informing on the value of personalizing gamification from a visual perspective, which has been scarcely explored in the literature [Hallifax et al. 2019].

Note that those insights should be interpreted in light of this paper's limitations. First, the study only focuses on badges and intrinsic motivation. While those are important aspects to consider, they might not capture the full complexity of gamification design and its impact on learning outcomes. Similarly, the study explored a single brand. Hence, future research could explore other game elements and brands, as well as additional learning outcomes, to have a more comprehensive understanding of how different design factors influence learning experiences. Second, the study is limited to 19 undergraduate students. Whereas this sample size might restrict the generalizability of the findings, it enabled us to yield promising, initial evidence for an underexplored field. Thereby, future studies should replicate the study with larger, more diverse samples to increase the robustness and validity of the results.

Furthermore, the study adopts a within-subject design. While this design might introduce order effects or carryover effects from one condition to another, we chose it as it enables a direct comparison within the same participants and used counterbalancing to

mitigate its issues. Additionally, the study is based on self-report measures. Although those are subjective and prone to biases, such as social desirability bias or participants' inability to accurately articulate their experiences, they often yield valid results based on past research. Nevertheless, we call for future research to explore objective measures, such as performance-based assessments or physiological measurements, as a way to test and ground our findings. To summarize, while this paper provides promising insights into the effects of badges' visual appearance on learning experiences, future research should consider and explore these limitations to further advance the understanding of gamification design and its impact on learning outcomes from a visual perspective.

6. Final Remarks

Gamification has been widely applied and studied within the educational domain. Whereas its overall effect on learning outcomes is positive, some empirical studies reveal cases wherein gamification is ineffective or harmful. Aiming to mitigate the latter cases, researchers are often concerned with improving the gamification design. In that regard, previous research has mainly focused on comparing the different game elements included in a design. On the other hand, there is promising evidence suggesting that properly selecting the game elements' aesthetics could benefit learning experiences. Despite that, there is a lack of empirical evidence from experimental studies on how students' experiences compare depending on whether game elements' visuals are aligned with their interests or not.

Based on that lack, our goal was to understand how thematic badges from a specific brand, affect learning experiences compared to non-thematic badges. For this, we conducted a mixed-methods, experimental study in which participants interacted with two versions of a gamified quiz: one used non-thematic badges (Medals/Trophies), while the other employed thematic badges (i.e., based on aesthetics from the Naruto brand). Overall, our findings suggested that i) students were more intrinsically motivated to interact with the Naruto-themed gamification, compared to the standard version, that ii) the more one's enthusiasm regarding the Naruto brand, the more motivating its badges were perceived to be compared to medals, and that iii) reasons for those results are enjoying the anime, not liking the medals, and having feelings for the brand.

Thus, this paper contributes to the gamification literature by providing empirical evidence on the impact of badges connected to students' feelings on learning experiences based on their aesthetics. Our results support the value of employing gamification designs that consider users' interests and make game elements visually appealing to enhance intrinsic motivation and engagement. Importantly, our insights also raise concern that such a connection might distract students and be ignored by those not interested in the visuals. Thus, our findings inform practitioners in designing effective gamified systems as well as guide future research in exploring the role of visual appearance in gamification design towards maximizing its contribution to learning outcomes.

Acknowledgments

We are deeply thankful to the students who participated in this study. This research was partially supported by the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie (grant No 101029543, GamInclusive).

References

Bai, S., Hew, K. F., and Huang, B. (2020). Is gamification "bullshit"? evidence from a meta-analysis and synthesis of qualitative data in educational contexts. *Educational Research Review*, page 100322.

Barbosa, S. D. J., Silva, B. d., Silveira, M. S., Gasparini, I., Darin, T., and Barbosa, G. D. J. (2021). Interação humano-computador e experiência do usuario. *Auto publicação*.

Cardoso, C. L., Gontijo, L. A., and Ono, M. M. (2017). Affective memory: An ethnographic approach to design. *Strategic Design Research Journal*, 10(1):79–88.

Chou, Y.-k. (2019). *Actionable gamification: Beyond points, badges, and leaderboards*. Packt Publishing Ltd.

Deterding, S., Dixon, D., Khaled, R., and Nacke, L. (2011). From game design elements to gamefulness: defining gamification. In *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments*, pages 9–15. ACM.

Hallifax, S., Serna, A., Marty, J.-C., and Lavoué, É. (2019). Adaptive gamification in education: A literature review of current trends and developments. In Scheffel, M., Broisin, J., Pammer-Schindler, V., Ioannou, A., and Schneider, J., editors, *Transforming Learning with Meaningful Technologies*, pages 294–307, Cham. Springer International Publishing.

Hanus, M. D. and Fox, J. (2015). Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. *Computers & Education*, 80:152 – 161.

Hicks, K., Gerling, K., Dickinson, P., and Vanden Abeele, V. (2019a). Juicy game design: Understanding the impact of visual embellishments on player experience. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, pages 185–197.

Hicks, K., Gerling, K., Richardson, G., Pike, T., Burman, O., and Dickinson, P. (2019b). Understanding the effects of gamification and juiciness on players. In *2019 IEEE Conference on Games (CoG)*, pages 1–8. IEEE.

Huang, R., Ritzhaupt, A. D., Sommer, M., Zhu, J., Stephen, A., Valle, N., Hampton, J., and Li, J. (2020). The impact of gamification in educational settings on student learning outcomes: a meta-analysis. *Educational Technology Research and Development*, pages 1–27.

Kao, D. and Harrell, D. F. (2018). The effects of badges and avatar identification on play and making in educational games. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–19.

Klock, A. C. T., Gasparini, I., Pimenta, M. S., and Hamari, J. (2020). Tailored gamification: A review of literature. *International Journal of Human-Computer Studies*, page 102495.

Koivisto, J. and Hamari, J. (2019). The rise of motivational information systems: A review of gamification research. *International Journal of Information Management*, 45:191–210.

Kotrlik, J. W. and Williams, H. A. (2003). The incorporation of effect size in information technology, learning, information technology, learning, and performance research and performance research. *Information Technology, Learning, and Performance Journal*, 21(1):1.

Linehan, C. and Kirman, B. (2017). Mc hammer presents: the hammer of transformative nostalgification-designing for engagement at scale. In *Proceedings of the 2017 CHI conference extended abstracts on human factors in computing systems*, pages 735–746.

Loughrey, K. and Broin, D. (2018). Are we having fun yet? misapplying motivation to gamification. In *2018 IEEE Games, Entertainment, Media Conference (GEM)*, pages 1–9. IEEE.

Mora, A., Riera, D., González, C., and Arnedo-Moreno, J. (2017). Gamification: a systematic review of design frameworks. *Journal of Computing in Higher Education*, 29:516–548.

Palomino, P. T., Toda, A. M., Rodrigues, L., Oliveira, W., Nacke, L., and Isotani, S. (2023). An ontology for modelling user'profiles and activities in gamified education. *Research and Practice in Technology Enhanced Learning*, 18.

Pedro, L. Z. (2016). *Uso de gamificação em ambientes virtuais de aprendizagem para reduzir o problema da externalização de comportamentos indesejáveis*. PhD thesis, Universidade de São Paulo.

Pereira, R., Reis, R., Oliveira, L., Derenievicz, G., Peres, L., and Silva, F. (2023). A liga do pensamento computacional: uma narrativa distópica para gamificar uma disciplina introdutória de computação. In *Anais do III Simpósio Brasileiro de Educação em Computação*, pages 205–215. SBC.

Pereira, R., Rodrigues, K. R., and Silveira, M. S. (2021). Gamifichi: thematized badges for hci courses. In *Proceedings of the XX Brazilian Symposium on Human Factors in Computing Systems*, pages 1–10.

Rodrigues, L., Arndt, D., Palomino, P., Toda, A., Klock, A. C. T., Avila-Santos, A., and Isotani, S. (2022a). Affective memory in gamified learning: A usability study. In *Anais do XXXIII Simpósio Brasileiro de Informática na Educação*, pages 585–596. SBC.

Rodrigues, L., Pereira, F., Toda, A., Palomino, P., Oliveira, W., Pessoa, M., Carvalho, L., Oliveira, D., Oliveira, E., Cristea, A., et al. (2022b). Are they learning or playing? moderator conditions of gamification's success in programming classrooms. *ACM Transactions on Computing Education (TOCE)*.

Rodrigues, L., Toda, A. M., Oliveira, W., Palomino, P. T., Avila-Santos, A. P., and Isotani, S. (2021). Gamification works, but how and to whom? an experimental study in the context of programming lessons. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*, pages 184–190.

Rodrigues, L., Toda, A. M., Palomino, P. T., Oliveira, W., and Isotani, S. (2020). Personalized gamification: A literature review of outcomes, experiments, and approaches. In *Eighth International Conference on Technological Ecosystems for Enhancing Multiculturality*, pages 699–706.

Ryan, R. M. and Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford Publications.

So, K. K. F., King, C., and Sparks, B. (2014). Customer engagement with tourism brands: Scale development and validation. *Journal of Hospitality & Tourism Research*, 38(3):304–329.

Toda, A. M., Klock, A. C., Oliveira, W., Palomino, P. T., Rodrigues, L., Shi, L., Bittencourt, I., Gasparini, I., Isotani, S., and Cristea, A. I. (2019). Analysing gamification elements in educational environments using an existing gamification taxonomy. *Smart Learning Environments*, 6(1):16.

Toda, A. M., Valle, P. H. D., and Isotani, S. (2018). The dark side of gamification: An overview of negative effects of gamification in education. In Cristea, A. I., Bittencourt, I. I., and Lima, F., editors, *Higher Education for All. From Challenges to Novel Technology-Enhanced Solutions*, pages 143–156, Cham. Springer International Publishing.

Tondello, G. F., Kappen, D. L., Ganaba, M., and Nacke, L. E. (2019). Gameful design heuristics: A gamification inspection tool. In *International Conference on Human-Computer Interaction*, pages 224–244. Springer.

Wilcox, R. (2017). *Introduction to robust estimation and hypothesis testing*. Elsevier, , 4th edition.

Wohlin, C., Runeson, P., Hst, M., Ohlsson, M. C., Regnell, B., and Wessln, A. (2012). *Experimentation in Software Engineering*. Springer Publishing Company, Incorporated.

Xi, N. and Hamari, J. (2019). The relationship between gamification, brand engagement and brand equity.