Towards Automating the Personalization of Gamified Learning's Aesthetics: A Survey Comparing Themed and Standard Badges Generated by AI

Luiz Rodrigues[1], Renan Vinicius Aranha[2], Newarney T. Costa[3], Lucas de Almeida Ribeiro[4] Heder Filho S. Santos[3], Cleon X. Pereira Júnior[3]

[1]Instituto de Computação - Universidade Federal de Alagoas
Maceió – AL - Brasil

[2]Faculdade de Engenharia - Universidade Federal de Mato Grosso
Cuiabá – MT – Brasil

[3]Campus Iporá - Instituto Federal Goiano (IF Goiano)
Iporá – GO - Brasil

[4]Campus Luziânia - Instituto Federal de Goiás (IFG)
Luziânia – GO – Brasil

luiz.rodrigues@nees.ufal.br, renan.aranha@ufmt.br

{cleon.junior, newarney.costa}@ifgoiano.edu.br

lucas.ribeiro@ifg.edu.br, heder.filho@estudante.ifgoiano.edu.br

Abstract. *Gamification has been effective in enhancing learning outcomes, though its impact varies due to individual learner differences, such as interests and personality traits. This paper presents a survey study (n = 107) examining opinions on themed versus standard badges automatically designed by Generative Artificial Intelligence (GAI), considering factors such as different themes, engagement with the theme, and personality traits. Preliminary results reveal that themed badges are generally more motivating than standard ones, regardless of the badge's theme or the personality traits. These findings suggest that educators and designers can explore GAI to explore themed badges without being constrained by theme-specific effectiveness.*

1. Introduction

Gamification provides a well-established impact on learning outcomes by increasing engagement, motivation, and retention through game-like elements such as points, badges, and leaderboards [Sailer and Homner 2020]. These elements can make learning more enjoyable and interactive, fostering a sense of achievement and competition among students [Bai et al. 2020]. However, the effectiveness of gamification varies significantly among learners due to differences in preferences, personality traits, and other factors [Van Roy and Zaman 2018, Rodrigues et al. 2020a, Tondello et al. 2017]. This variability highlights the need for personalized gamification designs that cater to individual learner needs and preferences [Klock et al. 2020].

Personalizing gamified experiences aims to ensure that all students can benefit from enhanced engagement and motivation, ultimately leading to more effective and in-

DOI: 10.5753/sbie.2024.244949

clusive learning outcomes [Xiao and Hew 2024]. It often involves tailoring game elements selection to individual learners' preferences, needs, and motivations to maximize engagement and learning effectiveness [Rodrigues et al. 2020b]. While this approach has great potential, it also presents limitations, such as increased complexity in design and implementation as well as the challenge of knowing which game elements to use for each learner [Xiao and Hew 2023]. On the other hand, an alternative yet under-explored approach is personalizing the game elements aesthetics, modifying aspects such visual and auditory elements [Hallifax et al. 2019].

Recent research has demonstrated increasingly interest on aesthetic-based personalization in gamified learning. In that line, [Linehan and Kirman 2017] backgrounds the relevance of game elements' aesthetics based on Pokémon Go's success compared to its counterpart Ingress, motivating subsequent research to address this issue within gamified learning context's. [Pereira et al. 2021], for instance, addresses a similar issue by gamifying a learning environment with badge aesthetics themed after renewed researchers from environment's learning subject. Accordingly, [Rodrigues et al. 2022, Rodrigues et al. 2023] presented laboratory studies aiming to understand how aesthetic-based personalization compares to the one-size-fits all approach. Similarly, [Pereira Júnior et al. 2023] compared the same approaches, but in an 4-week experimental study in the context of high-school lessons. In common, all of those studies provided promising evidence that aesthetic-based personalization might contribute to learning outcomes compared to the one-size-fits-all approach.

Notably, related work has explored different themes to personalize the game elements' aesthetics, and all of them concerned the *acknowledgment* game element [Toda et al. 2019]. [Pereira et al. 2021] used renewed researchers from the subject being taught to personalize badges' aesthetics. [Rodrigues et al. 2023] and [Rodrigues et al. 2022] used Naruto and Pokémon as themes, respectively, to tailor badges. [Pereira Júnior et al. 2023] performed an a priori investigation with their target students to understand and use their preferences to create themed, personalized badges. In common, none of these studies investigated how different themes compare to each other, and whether different themes affect how aesthetic-based personalization compares to the one-size-fits-all approach. Consequently, an open Research Question (RQ) is:

> **RQ1:** *Does the advantage of aesthetic-based personalized badges over standard ones depend on the badge's theme/brand?*

Another consideration is that [Rodrigues et al. 2023] found that the more one's engagement with the theme (Naruto in that case), the more they felt motivated by the themed badges compared to the standard ones. This aligns with research demonstrating that people's individual factors, such as personality traits, affect how they perceive and are more by gamification. Nevertheless, related work on aesthetic-based personalization did not address this issue. Thus, another RQ is:

> **RQ2:** *How do people's perceptions regarding themed and standard badges compare depending on their engagement with the theme and personality traits?*

Furthermore, another important issue is the feasibility of creating aesthetically relevant badges. One known issue in gamified learning is the challenges associate with

creating compelling designs [Xiao and Hew 2023, Klock et al. 2020]. Accordingly, deploying aesthetic-based personalized gamification also suffers from this challenge, as one must create multiple game elements with varied aesthetics. The rapid rise of generative Artificial Intelligence (AI) presents an alternative to alleviate this issue, enabling gamification designers to automate the generation of aesthetically relevant themed badges. However, an open RQ in that regard is:

> **RQ3:** *How does themed badges generated by AI compare to badges concerning the one-size-fits-all approach?*

This paper presents an empirical study towards addressing those RQs. For this, we are conducting a survey-based study to understand people's opinions regarding AI-generated badges compared to a standard badge in terms of different themes. Furthermore, acknowledging that people's intrinsic differences affect their perceptions, we are investigating how those perceptions differ depending on people's engagement with each theme as well as their personality traits. Therefore, this ongoing research differs from past studies by i) using AI to generate personalized badges, ii) comparing different brands to one another based on the same sample of participants, and iii) considering how personality traits, besides one's engagement with the brand, affect people's perceptions. Thus, this paper provides initial evidence on whether themed badges offer an advantage over standard ones, how people's engagement with the theme and their personality traits influence this advantage, and whether generative AI can create badges that people find acceptable.

2. Method

This research used a survey-based study to answer its RQs. Surveys allow researchers to gather data from a large, diverse population quickly and cost-effectively, providing a broad understanding of trends and patterns [Rea and Parker 2014]. In contrast, case/experimental studies can be expensive and resource-intensive, often requiring extensive planning and execution, beside carrying higher risks, such as potential biases or ethical concerns [Wohlin et al. 2012]. Surveys can help identify areas of interest and inform more targeted, detailed studies, minimizing the risk of costly mistakes, making them an ideal tool for exploratory research and allowing for efficient hypothesis generation and preliminary insights [Creswell and Creswell 2017]. Therefore, considering aesthetic-based personalization is a recent research line [Pereira Júnior et al. 2023, Rodrigues et al. 2023], we considered a survey to be a suitable approach for this work in progress.

This study's procedure consists in three steps. The first step was **survey development**, which led to a survey composed of four sections. The *first section* introduces the research, asking for participants to provide informed consent that they agree with using their data for research purposes. The *second section* collected the participant's interest regarding the themes used in the badges' aesthetics (see details next). For this, we followed related work and collected answers through a seven-point Likert-scale and the following statement, extracted from the Emotional Engagement Scale *I feel enthusiastic about* <*THEME*> [Rodrigues et al. 2023]. Here we opted for a single item to reduce the effort required to complete the survey and, hence, increase its validity [Wohlin et al. 2012].

The *third section* collected the participant's opinions regarding different badges. For increased reliability, we asked the participant to consider the following scenario: *You*

Figure 1. Example of badges displayed to participants. On the left, the standard badge. On the right, the Game of Thrones themed badge.

are using a Virtual Learning Environment called Moodle to help you with your studies. After each on-time submission, Moodle rewards students with badges to motivate them to submit more. Next, *fourth section* presented side-by-side images, where one was a standard badge (i.e., a medal) and the other was an AI-generated badge of one of the following brands: Super Mario, Game of Thrones, Naruto, and Avengers (see Figure 1). Using a medal as the standard badge is based on related work (e.g., [Pereira et al. 2021, Rodrigues et al. 2022, Pereira Júnior et al. 2023]), whereas the selected brands are based on data collected in a prior study with a similar sample [Pereira Júnior et al. 2023]. Importantly, all badges were generated using the same AI (i.e., DALL·E 2[1]) and the same prompt (i.e., "<*Theme*> themed medal.", where <*Theme*> has been replaced by: Super Mario Bros, Game of Thrones, Naruto or Avengers).

For each side-by-side image in the survey's section three, it presented a forced-choice item. That is, participants should indicate whether, based on the described scenario, they would feel more motivated by the standard badge, the themed (i.e., brand-based) badge, neither of them, or equally motivated by both. This method enhances survey validity by compelling respondents to make explicit choices, thus yielding clearer insights into their motivations, and reduces the ambiguity and social desirability bias often associated with traditional Likert scales, as it eliminates the neutral option and forces a decision [Orbán-Mihálykó et al. 2024, Ramík and Ramík 2020]. Additionally, it optimizes the survey by focusing on direct comparisons rather than evaluating every item independently, thereby, reducing the number of survey items, minimizing respondent fatigue, and enhancing the quality and accuracy of the data collected [Wohlin et al. 2012].

Next, the survey's *last section* was participant characterization. Besides capturing demographic information (i.e., age, gender and living region), it captured the participant's personality traits. The study of human personality through traits can enhance, for example, the comprehension of their behaviors and characteristics, such as interest in exploring new resources and artistic appeal, in a diverse setting, not limited to a specific context of use, such as digital games [Ismail and Soja 2023, Chen et al. 2023]. Furthermore, understanding how personality traits relate to badge preferences can support the development of personalized solutions for creating these resources tailored to each student. Then, the Big Five Factors of Personality was chosen to analyze the volunteers' personalities in this study. This model, proposed by [Costa and McCrae 1999], organizes human personality into five factors: agreeableness, conscientiousness, extraversion, neuroticism and open-

[1]https://openai.com/index/dall-e-2/

ness to experience.

The second step was **data collection**. For this, the authors disclosed the survey link through social media and email lists, aiming for a large yet reliable sample of students from Brazilian technical or higher education institutions. Currently, we have received 141 answers, of which 34 had to be discarded based on reliability checks, such as out-of-target-sample (underage people), unreasonable genders (e.g., *retroescavadeira*), and unacceptable answer times (e.g., taking 60 seconds to complete the survey). Based on that, our sample consists of 107 participants (104 from Brazil's Midwest region), with an average age of 23 years (ranging from 18 to 44), wherein 80 of them identify as males, 26 as females, and one preferred not to answer it.

The third step was **data analysis**. According to the exploratory nature of our RQs, we followed literature recommendations and mainly analyzed the survey answers with data visualization techniques, such as barplots and scatterplots [Vornhagen et al. 2020, Cairns 2019]. Nevertheless, we also conducted Chi-squared homogeneity and independence tests [Wohlin et al. 2012] to grasp an initial understanding of how the participants opinions vary depending on the badge design and the brand. Additionally, we also use data visualization techniques to understand how participants perceptions differ depending on their engagement with the brand, as well as their personality traits. As a work in progress, however, we limit this analysis to investigate how the percentage favoring the themed badges changes depending on these personal factors, such as engagement with the theme, increase due to the sample size.

3. Preliminary Results

This section presents the preliminary results from this work in progress, which features 428 registers (four for each of the 107 answers, given that the survey concerned badges from four themes) at the time of writing. Table 1 presents the participants' overall opinions regarding the different badge designs. It demonstrates that the themed badges were substantially more favored (57%) than the standard ones (24%), with less than 20% of the answers choosing neither or both designs. Accordingly, the Chi-squared test of homogeneity demonstrates that this difference is statistically significant (χ = 253.40, p-value = < 0.001). Hence, this finding provides empirical evidence that themed badges were considered to be more motivating than standard, medal-based ones even when AI generated.

Table 1. Participants overall opinions regarding badge designs.

	Themed	**Standard**	**Neither**	**Both**
Count (%)	243 (57%)	102 (24%)	44 (10%)	39 (9%)

Next, we investigated whether the themed badges predominance differs depending on their theme (i.e., the brand they are based on). Figure 2 demonstrates how participants' opinions varies depending on both badge design and the brand used to personalize them, supporting the finding that themed ones are considered more motivating than the standard ones. Accordingly, the Chi-squared test of homogeneity demonstrates there is no dependence between badge design and brands (chi = 9.09, p-value = 0.43, degrees of freedom = 9). Thereby, this finding provides empirical evidence that the extent to which participants feel more motivated by themed badges does not differ significantly depending on the brand they are based on.

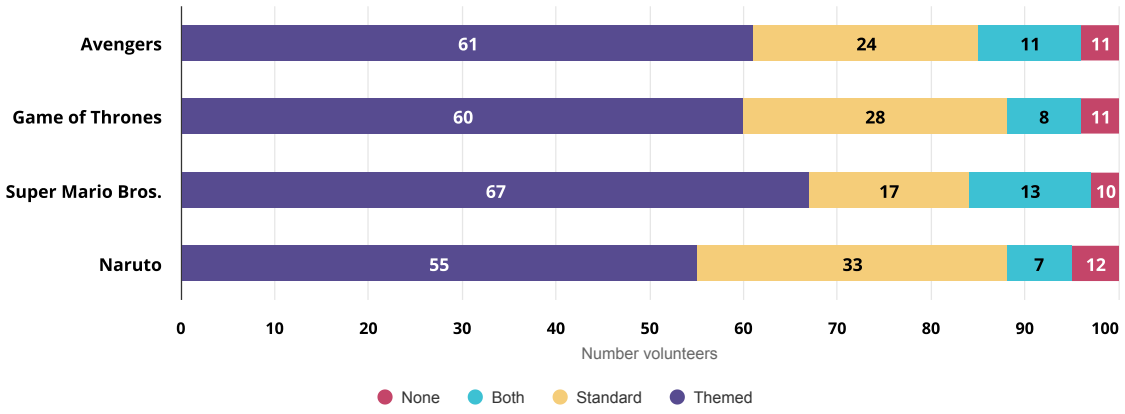

Figure 2. Participants overall opinions regarding badge designs by theme/brand.

Finally, we investigated whether the participants' opinions differed depending on their personal aspects. Considering brand did not affect the results (see Figure 2), this analysis focused on whether the percentage of people that preferred themed badges changed depending on their engagement with the brand and their personality traits. Table 2 demonstrates that the percentage of participants that prefer themed badges increases for those with higher overall engagement with the brand. Also, Figure 3 demonstrates even when grouping the volunteers according to their predominant personality trait, the results suggest that there is no clear differences in people's opinions for themed badges. In all cases, this option was the one that led to the highest engagement levels among the volunteers. In contrast, a more specific analysis indicates personality traits possibly play a role when comparing people's perceptions of badge design depending on the badge's brand (see Figure 4).

Table 2. People who chose themed badges by level of engagement with brand.

Engagement Level	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Those who chose themed badges	44%	41%	57%	43%	57%	63%	72%

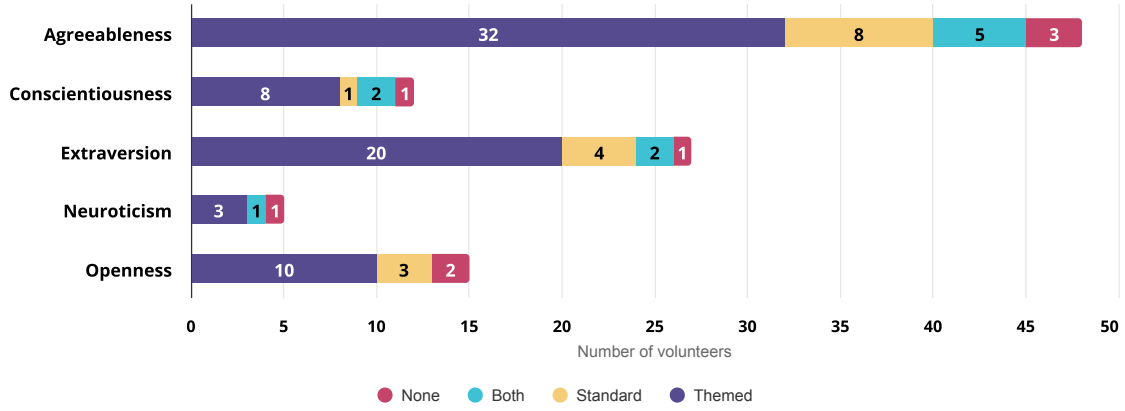

Figure 3. Distribution of participants' preferences regarding the type of badges according to their predominant personality trait.

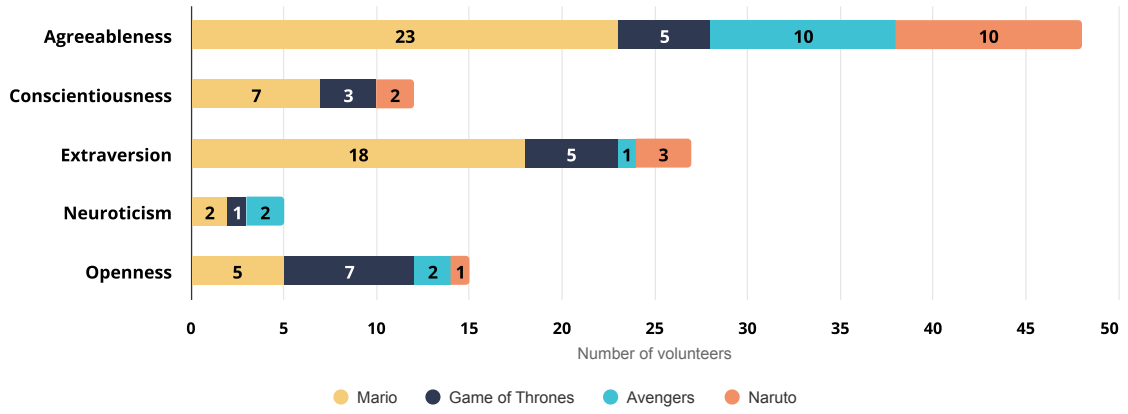

Figure 4. Distribution of participants' preferences regarding the theme of badges according to their predominant personality trait.

4. Discussion and Next Steps

The preliminary findings of this work present encouraging evidence regarding the potential of themed badges to motivate students to engage with educational tasks. The results suggest that people see themed badges as substantially more motivating than standard ones (see Table 1), and that this difference does not depend on the badge's theme (see Figure 2) or the personality traits (see Figure 3), since the preference for themed badges remained consistent across all user groups. On the other hand, our findings suggests that the personality traits can slightly influence the preference for a theme/brand (see Figure 4). However, we highlight more data is necessary to draw definitive conclusions, as the traits involved (neuroticism and openness) had a smaller number of participants.

Mainly, these findings support the value from exploring aesthetic-based personalization in gamified learning. Considering that the motivating effect of themed badges is not dependent on the specific theme chosen, this finding suggests that the appeal of themed badges may be related to their novelty or uniqueness, rather than any particular theme itself. Therefore, designer and educators could potentially utilize a variety of themes to maintain student interest and motivation without worrying about the effectiveness being limited to certain themes.

Complementary, it is important to note that the data suggests that the level of engagement students have with a badge's theme is positively associate with the motivational impact of themed badges. This suggests that when students have a pre-existing interest or connection with the brand associated with the badge, they are even more likely to find the themed badges motivating. This finding highlights the importance of considering students' interests and preferences when designing gamified educational elements to maximize their motivational impact. Thus, these findings provide encouraging evidence that personalized gamification, through the use of themed badges, holds promise as an effective strategy for increasing student motivation and engagement in educational settings.

These insights corroborate recent research that investigated the role of themed badges by personalizing them to students' interests in laboratory [Rodrigues et al. 2023] and field [Pereira Júnior et al. 2023, Pereira et al. 2021] studies. However, previous studies are limited to investigating how a single theme compares to standard badges. More-

over, they depend on a manual process to generate the aesthetically-personalized badges, which is laborious and likely to complicate deploying such kind of personalization. Finally, they did not consider how individual aspects, such as personality traits, affect people's perceptions of how aesthetically-personalized and standard badges compare. Thus, this research expands the literature by investigating how badges of different themes compared to a standard one, within the same sample, besides addressing the potential of generative AI for themed badge generation as well as seeking to understand how people's individual aspects influence their opinions.

Nevertheless, we acknowledge this is an ongoing research that needs further investigation to ground its insights. Therefore, we plan a number of future steps to follow up with this study. Our main aim is to increase this survey's sample. While we have an interesting number of answers (i.e., 107), we would like to increase our dataset's representativeness, for instance by including answers from people of other Brazilian regions. Accordingly, we plan to apply more advanced data analysis techniques to understand our data. For example, the different ways in which each personality trait affects people's opinions suggest that they should be analyzed together, considering their interactions, rather than independently. Hence, with an increased sample size, we expect to be able to reliably perform more advanced data analysis techniques.

Finally, we plan to design and execute empirical studies comparing themed and standard badges. Aiming to understand how our survey results apply to practice, we plan to start with experimental studies and, depending on their outcomes, evolve for field studies. Our ultimate goal is to ground the understanding of how aesthetic-based personalization of gamification affects learning outcomes. For this, we plan to follow the design of related work (e.g., [Rodrigues et al. 2023, Pereira Júnior et al. 2023]) for comparability, informed by relevant experimental studies guidelines [Wohlin et al. 2012].

5. Final Remarks

Personalization has been widely investigated to maximize gamification's effects on learning outcomes. While studies have mostly focused on tailoring the game elements availability, yielding inconclusive results on whether it improves the one-size-fits-all approach, recent research has highlighted the potential of personalizing the aesthetics of game elements. However, related work has not compared how different themes affect aesthetic-based personalization compared to the one-size-fits-all approach or generative AI's potential to optimize the generation of themed badges for personalized gamification designs. To address that lack, this paper presents an ongoing research based on a survey.

Our preliminary findings indicate that themed badges are considered significantly more motivating than standard ones, irrespective of the badge's theme. Importantly, engagement with the badge's theme further enhances this motivational potential, suggesting the importance of personal interests in designing gamified educational elements. This research supports aesthetic-based personalization in gamified learning and proposes that novelty or uniqueness, rather than theme specificity, drives the appeal of themed badges. Therefore, this work in progress expands on previous research by comparing multiple themes within the same sample and exploring generative AI for badge creation. Accordingly, our future plans include increasing the survey sample size and conducting empirical studies to validate the findings in practical settings.

6. Artifacts Availability

The artifacts generated from this study are available from the corresponding author

7. Acknowledgments

Funding-related acknowledgements removed for review. We acknowledge the use of generative artificial intelligence tools, such as ChatGPT (3.5), Grammarly, and Google Translate, to aid in writing and revising this paper. The authors conducted a thorough review of the text and assume full responsibility for its content.

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.

Cairns, P. (2019). *Doing better statistics in human-computer interaction*. Cambridge University Press.

Chen, Q., Christensen, A. P., Kenett, Y. N., Ren, Z., Condon, D., Bilder, R., Qiu, J., and Beaty, R. (2023). Mapping the creative personality: A psychometric network analysis of highly creative artists and scientists. *Creativity Research Journal*, 35:455 – 470.

Costa, P. and McCrae, R. (1999). A five-factor theory of personality. *Handbook of personality: Theory and research*, 2(01).

Creswell, J. W. and Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

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.

Ismail, T. and Soja, J. (2023). Relationship between personality traits and creative thinking among secondary school students. *i-manager's Journal on Educational Psychology*.

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.

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.

Orbán-Mihálykó, É., Mihálykó, C., and Gyarmati, L. (2024). Evaluating the capacity of paired comparison methods to aggregate rankings of separate groups. *Central European journal of operations research*, 32(1):109–129.

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.

Pereira Júnior, C. X., Heder Filho, S. S., Rodrigues, L., and Costa, N. T. (2023). Investigating the effectiveness of personalized gamification in enhancing student intrinsic motivation: an experimental study in real context. In *Anais do XXXIV Simpósio Brasileiro de Informática na Educação*, pages 838–850. SBC.

Ramík, J. and Ramík, J. (2020). Pairwise comparison matrices in decision-making. *Pairwise comparisons method: Theory and applications in decision making*, pages 17–65.

Rea, L. M. and Parker, R. A. (2014). *Designing and conducting survey research: A comprehensive guide*. John Wiley & Sons.

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

Rodrigues, L., Toda, A. M., Oliveira, W., Palomino, P. T., and Isotani, S. (2020a). Just beat it: Exploring the influences of competition and task-related factors in gamified learning environments. In *Anais do XXXI Simpósio Brasileiro de Informática na Educação*, pages 461–470. SBC.

Rodrigues, L., Toda, A. M., Palomino, P. T., Klock, A. C., and Avila-Santos, A. (2023). Motivation no jutsu: Exploring the power of badge aesthetics in gamified learning. In *Anais do XXXIV Simpósio Brasileiro de Informática na Educação*, pages 631–643. SBC.

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

Sailer, M. and Homner, L. (2020). The gamification of learning: a meta-analysis. *Educ Psychol Rev*, 32:77–112.

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.

Tondello, G. F., Mora, A., and Nacke, L. E. (2017). Elements of gameful design emerging from user preferences. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, pages 129–142. ACM.

Van Roy, R. and Zaman, B. (2018). Need-supporting gamification in education: An assessment of motivational effects over time. *Computers & Education*, 127:283–297.

Vornhagen, J. B., Tyack, A., and Mekler, E. D. (2020). Statistical significance testing at chi play: Challenges and opportunities for more transparency. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, pages 4–18.

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

Xiao, Y. and Hew, K. F. (2023). A systematic literature review on personalized gamification: Algorithms and techniques. *EdMedia+ Innovate Learning*, pages 1318–1325.

Xiao, Y. and Hew, K. F. (2024). Personalised gamification enhances student participation but produces mixed effects on emotional and cognitive engagements: a systematic review. *Interactive Learning Environments*, pages 1–27.