

Usability Evaluation of an LMS-notion Integrated Badge-Based Learning System

Beyza Altuner¹, Muhtar Çağkan Uludağlı¹ and Esin Çağlayan²

¹Department of Computer Engineering, Faculty of Engineering, Izmir University of Economics, Türkiye

²Teaching and Learning Center, Izmir University of Economics, Türkiye

beyza.altuner@ieu.edu.tr

cagkan.uludagli@ieu.edu.tr

esin.caglayan@ieu.edu.tr

Abstract: This study evaluates the perceived usability of a badge dashboard implemented as a two-platform architecture in an undergraduate microprocessors course. Digital badges are increasingly adopted in higher education as mechanisms for representing competencies and motivating learner progress. However, less attention has been given to the usability of the dashboard interfaces through which students access and interpret badge-related information, particularly when these systems span multiple platforms. A single-cohort post-use usability evaluation was conducted following one academic semester during which 52 students interacted with a digital badge catalog and leaderboard embedded into the Blackboard Learning Management System (LMS) via Notion pages. The study focused specifically on the usability of the integrated Blackboard-Notion badge dashboard through which students accessed badge-related information, rather than on the educational effectiveness of badges or gamification itself. Usability was assessed using the System Usability Scale (SUS) on a screened sample of 47 respondents (N = 47). The system achieved a mean SUS score of 60.11, corresponding to an “OK” adjective rating within the marginal acceptability range. Item-level patterns revealed relatively strong learnability and a low perceived need for technical assistance, indicating that the system supports independent operation without substantial onboarding. In contrast, ratings for overall usage preference, confidence, ease of use, and perceived integration were more inconsistent; suggesting that while the workflow is learnable, it is not consistently experienced as seamless. This pattern is consistent with the interaction costs that arise when an institutional LMS embeds a third-party interface with distinct conventions, creating navigational friction that can undermine perceived coherence even when individual components function adequately. Exploratory subgroup analyses found no statistically significant differences in usability by gender or prior badge-system familiarity, and no significant association between academic performance and SUS scores. These findings highlight opportunities to improve the perceived integration and information architecture of multi-platform badge dashboards while demonstrating that such systems can remain learnable in authentic LMS settings.

Keywords: Digital badges, Gamification, Badge-based learning (BBL), Learning management system (LMS), System usability scale (SUS)

1. Introduction

Digital badges are increasingly used in higher education to represent competencies (Shields and Chugh, 2017), validated accomplishments (Elliott, Clayton and Iwata, 2014), and assessment-linked achievements (Newby and Cheng, 2020). In many institutional settings, badge-based learning (BBL) systems are deployed within Learning Management Systems (LMS) or are extended through third-party tools that offer richer catalogs, dashboards, or social comparison (Christy and Fox, 2014). While the pedagogical value of badges is widely discussed, empirical evidence regarding their effectiveness remains limited (Dicheva et al, 2015). From a human-computer interaction (HCI) perspective, badges and leaderboards are not merely incentive mechanisms; they are interaction artifacts that shape how learners perceive progress, interpret requirements, and navigate course workflows. When badge-related information is distributed across an LMS and an external dashboard, students must manage additional navigation steps, distinct interface conventions, and potentially inconsistent mental models. Such fragmentation may increase cognitive load, reduce perceived integration, and introduce interaction friction that can undermine system quality and sustained use, even when the underlying badge logic is clear.

In response to these usability and integration challenges, this paper presents and evaluates a pragmatic, replicable badge dashboard architecture in which the institutional LMS (Blackboard) functions as the primary access layer, while Notion provides an embedded, database-driven visualization layer for a badge catalog and a live leaderboard. The system was implemented in an undergraduate assembly language programming course, where well-documented instructional challenges make it a suitable context for deploying badge-based learning systems (Uludağlı, 2025). The primary objective is to document the system and evaluate its perceived usability under authentic classroom conditions using the System Usability Scale (SUS), developed by Brooke (1996).

Accordingly, the study is guided by the following research questions:

RQ1: What is the overall perceived usability of the integrated Blackboard-Notion badge dashboard, as measured by the SUS?

RQ2: Which aspects of usability (as reflected in SUS item-level patterns) constitute relative strengths and weaknesses of the system?

RQ3: Do perceived usability outcomes differ as a function of learner characteristics and prior experience?

The contributions of this paper are threefold. First, it documents a lightweight integration of a third-party dashboard (Notion) within Blackboard while maintaining a separation between grading data and student-facing progress visualizations. Second, it provides empirical usability evidence for this integration using the System Usability Scale (SUS). Third, it derives design implications for instructors and designers seeking to deploy badge dashboards that balance transparency and motivational affordances with low-friction interaction.

2. Background

Gibson et al (2015) describe digital badges as visible representations of competencies and progression within educational technology environments. Deterding et al (2011) define gamification as the use of game design elements in non-game contexts, positioning badges as a recurring game interface design pattern that can be repurposed for educational and other non-game environments. The motivational functions of badges have been widely examined using Self-Determination Theory as a conceptual framework (Ryan and Deci, 2000), reviewed across multiple empirical studies (Hamari, Koivisto and Sarsa, 2014), and incorporated into broader gamification designs (Brambilla et al, 2024). However, less attention has been directed toward the usability characteristics of badge delivery systems themselves.

The integration of badge systems with institutional LMS platforms presents both opportunities and challenges. While LMS environments provide stable access and centralized authentication, many mainstream implementations offer limited support for rich badge catalogs or dynamic leaderboards. These constraints have motivated hybrid approaches in which third-party tools are embedded within LMS course shells (Sousa-Vieira, Ferrero-Castro and López-Ardao, 2022). When badge information is distributed across an LMS and an embedded external dashboard, learners must manage additional navigation steps, interpret differing interface conventions, and maintain coherent mental models across environments. Such fragmentation can create interaction friction despite pedagogically sound badge designs.

Empirical research on digital badges in education has primarily focused on motivational and performance outcomes rather than interface usability. Prior studies have examined motivation, learning performance (Hanus and Fox, 2015), badge integration (Sousa-Vieira, Ferrero-Castro and López-Ardao, 2022), and students' perceptions of digital badges (Gregg et al, 2022). However, relatively few studies have employed standardized usability instruments to evaluate the interface characteristics of badge delivery systems themselves. Bangor, Kortum and Miller (2009) demonstrated that the System Usability Scale (SUS) provides a reliable measure of perceived usability across a wide range of interactive systems. Yet its application to multi-platform digital badge systems remains underexplored. This represents a notable gap, as effective dashboard usability is essential for enabling the sustained engagement and transparent progress visualization that badge systems aim to promote.

Design guidance for badge systems emphasizes aligning badge criteria with learning objectives (Park and Kim, 2019) and providing information-rich metadata to support transparent credentialing (Casilli and Hickey, 2016). Complementary work on gamification design also highlights the importance of fostering meaningful user engagement (Deterding, 2012). However, limited guidance exists on the interaction-level usability challenges learners encounter when using such systems.

Taken together, the literature indicates a persistent gap between the widespread adoption of BBL designs and systematic evaluations of their interaction quality when implemented across multiple platforms. While prior work has clarified motivational mechanisms and pedagogical alignment, little is known about how learners experience the day-to-day usability of embedded badge dashboards that span an institutional LMS and an external visualization layer. Questions remain regarding the usability of badge dashboards spanning institutional LMS platforms and external visualization layers in authentic classroom settings. Addressing these issues requires empirical usability evidence grounded in real instructional deployments and assessed with standardized instruments such as the SUS.

3. Methodology

3.1 Digital Badge Design and Awarding Criteria

The badge set was produced through an iterative, prompt-driven visual synthesis process using a text-to-image generative artificial intelligence model. For each badge, the design specification constrained (i) the learning outcome, course concept, or competency represented; (ii) a metaphor aligned with the course domain such as role-playing game motifs to communicate “quests” or “boss” milestones; and (iii) shared visual rules (fixed aspect ratio, controlled palette, and recurring frame geometries). These constraints minimized visual variation that could affect badge recognition and dashboard scanning. The structured prompt used to generate the badges is presented in Figure 1a, while the resulting badge set and their associated acquisition criteria are shown in Figure 1b.

Badge awarding criteria were designed using performance-based thresholds derived from historical course data. Specifically, the average grades of all assessments of students from the previous academic year were used as benchmark values for each badge. Students in the current cohort were awarded a badge when their corresponding grade met or exceeded these predefined thresholds, ensuring that badge acquisition reflected relative achievement against a historical course baseline.

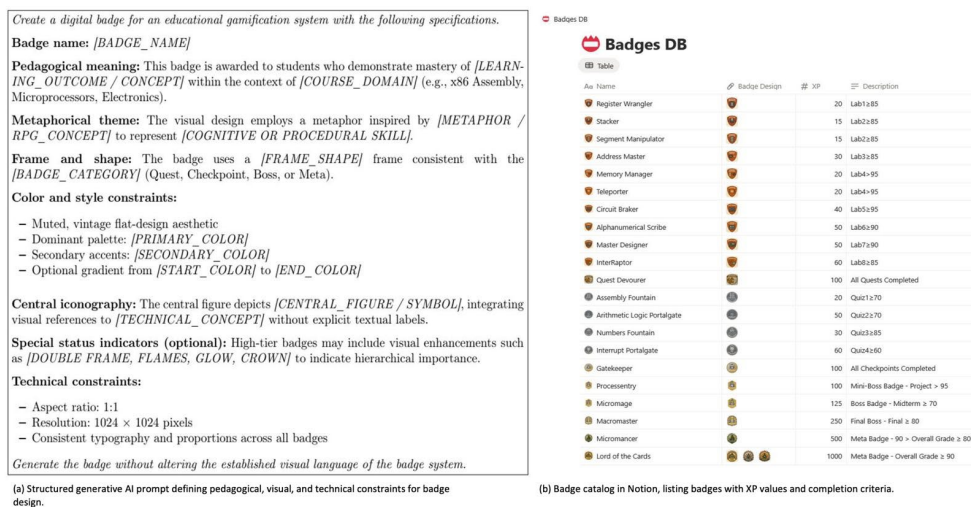


Figure 1: Badge system design pipeline with the structured generative design prompt to generate the badge catalog integrated into the learning environment

3.2 System Architecture

The BBL system was implemented as a lightweight two-platform architecture combining the Blackboard LMS (primary delivery and access layer) with Notion (embedded dashboard layer). Blackboard hosted the official course shell (CE342 - Fundamentals of Microprocessors) and served as the central navigation environment under institutional authentication. Within Blackboard, a dedicated *Gamification* area was created to provide a single-entry point for the badge catalog, leaderboard, and an overview of the gamification procedure (shown in Figure 2a). Notion content was embedded into Blackboard content items as read-only *iframe* views. Two embedded pages operationalized the dashboards. The first was a BBL landing page (Figure 2b) containing introductory information and an embedded badge catalog (Figure 1b), where students could review available badges, XP values, and acquisition criteria. This view served as an always-available reference model indicating what could be earned and how badge criteria related to course activities.

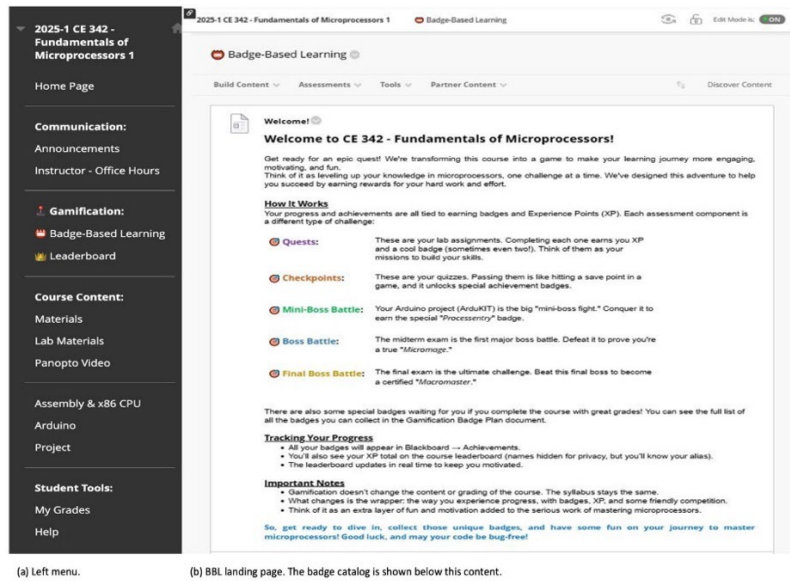


Figure 2: Blackboard course interface for the BBL system, showing the gamification-specific navigation entry (left) and the corresponding BBL landing page (right), where the Notion-hosted badge catalog is embedded and presented to students as part of the LMS environment

Second, a separate Blackboard item titled *Leaderboard* embedded a Notion database view (*Leaderboard DB*) that visualized cumulative progress. The leaderboard displayed student numbers, (i) an *Achievements* field populated with earned badges, and (ii) a *Total XP Points* field, enabling rank ordering (Figure 3).

To preserve privacy, the system separated instructor-only and student-facing data. A private Notion workspace contained assessment records, badge-awarding logic, and score calculations, while students accessed only a public linked view displaying anonymized identifiers, earned badges, and cumulative XP.

Although the student-facing Notion view was implemented using Notion's public sharing mechanism to enable embedding, it was not published or promoted as a publicly accessible webpage. The dashboard displayed only badge-related information (student identifier, earned badges, and accumulated XP), while the underlying assessment records, badge awarding logic, and course scores remained in a private instructor-only Notion workspace. No assessment grades or additional personal information were exposed through the dashboard.

Student Number	Achievements	Total XP Points
S 01	Register Stranger, Assembly Foundation, Address Master, Responder, Memory Manager, Circuit Breaker, Arithmetic Logic Portagale, Microimage, Alphanumeric Scribble, Numbers Foundation - 7	1025
S 02	Register Stranger, Assembly Foundation, Address Master, Circuit Breaker, Arithmetic Logic Portagale, Microimage, Responder, Master Designer, Interpolator, Memory Manager - 3	1005
S 03	Register Stranger, Segment Manipulator, Stack, Address Master, Circuit Breaker, Microimage, Alphanumeric Scribble, Numbers Foundation, Master Designer, Interpolator - 4	1045
S 04	Register Stranger, Assembly Foundation, Address Master, Circuit Breaker, Microimage, Alphanumeric Scribble, Numbers Foundation, Master Designer, Interpolator, Internet Portagale - 3	1035
S 05	Register Stranger, Stack, Segment Manipulator, Address Master, Memory Manager, Responder, Circuit Breaker, Microimage, Alphanumeric Scribble, Numbers Foundation - 4	1075

Figure 3: Leaderboard view embedded in Blackboard via Notion. The figure displays the final leaderboard of the course, with earned badges and corresponding XP points for the top five students using anonymized student identifiers. Remaining entries continue below in the LMS interface

From the student perspective, the workflow consisted of accessing the Blackboard course, reviewing the embedded badge catalog, and monitoring progress through the leaderboard. This design aimed to minimize context switching by keeping access within Blackboard while leveraging Notion's database affordances for dense, scrollable tabular visualization and straightforward web-publication options.

3.3 Study Design

This study employed a single-cohort, post-use usability evaluation of the integrated badge dashboard. Data were collected in Week 13 and Week 14 of the semester, after students had repeated exposure to the badge catalog and leaderboard, at a time when only three of the twenty-one total badges had yet to be awarded. The study protocol received institutional ethics approval. Participation was voluntary and based on informed consent. Students were informed that their decision to participate or not participate would have no influence on course grades or academic evaluation, thereby reducing the potential for instructor-student coercion.

Throughout the semester, students interacted with multiple gamification components, including digital badges, XP points, the badge catalog, the leaderboard, and performance-based badge thresholds. However, the post-study questionnaire explicitly instructed participants to evaluate only the usability of the integrated Blackboard-Notion badge dashboard through which this information was presented, rather than the pedagogical effectiveness or motivational value of these gamification elements.

3.3.1 Measurement instruments

Perceived usability was assessed using the SUS (Brooke, 1996), a 10-item instrument that provides a global measure of perceived usability. The unit of evaluation in this study was the student-facing integrated Blackboard-Notion badge dashboard, including its navigation, badge catalog, leaderboard, and presentation of badge-related information. Prior to completing the questionnaire, participants were informed that the SUS evaluated the combined badge achievement system consisting of the Blackboard course management system and the embedded Notion database used to display badges and the leaderboard. Accordingly, participants were asked to evaluate the usability of this interface rather than the educational effectiveness of badges, XP values, performance thresholds, or the broader gamification strategy. The original English version of the SUS was administered without translation. To direct participants toward the intended unit of evaluation, the generic term *system* in the SUS was adapted where appropriate to *badge achievement system*, while preserving the original response format and scoring procedure. Consequently, the SUS responses reflect perceived interface usability and interaction quality rather than attitudes toward the gamification design or its motivational effects. Items were rated on a 5-point Likert scale (strongly disagree to strongly agree) with alternating valence. Following standard scoring procedures, positively worded items were transformed using $(x - 1)$ and negatively worded items using $(5 - x)$. Transformed values were summed and multiplied by 2.5 to yield a 0 – 100 score. Consistent with established practice, the SUS was treated primarily as an overall indicator of usability (Lewis and Sauro, 2009; Bangor, Kortum and Miller, 2008; Bangor, Kortum and Miller, 2009), as individual items offer only limited diagnostic value. This approach aligns with recommendations cautioning against item-level interpretation, given evidence that single items are not meaningful when examined in isolation (Bangor, Kortum and Miller, 2008). Additionally, item-level responses were examined descriptively to highlight recurrent usability strengths and potential sources of interaction friction.

3.3.2 Participants and data quality screening

The course cohort consisted of 60 enrolled undergraduate students. Three students received non-completion outcomes (NA or W), did not participate in any of the gamified activities and were excluded from eligibility, yielding 57 eligible students. Of these, 52 students completed the usability questionnaire. Participants were predominantly within the 21 – 23 age range ($n = 38, 73.1\%$), followed by those aged 18 – 20 ($n = 9, 17.3\%$), with a smaller proportion older than 23 years ($n = 5, 9.6\%$). General academic performance, measured by self-reported cumulative grade point average (GPA), showed a mean of 2.41 ($SD = 0.46$) on a 4.00 scale. Most students were taking the course for the first time ($n = 39, 75\%$), while a smaller subset had repeated the course ($n = 13, 25.0\%$). Among the final analytical sample ($N = 47$), 22 participants (46.8%) reported prior familiarity with digital badge or micro-credential systems, while 25 (53.2%) reported no prior familiarity.

To mitigate the influence of careless responding, five cases were excluded: two due to strict straight-lining (all 5s or all 1s) and three due to implausibly low within-respondent variability across the alternating-valence SUS items (i.e., response patterns showing minimal dispersion inconsistent with attentive responding). These exclusions reduced the final analytic sample to $N = 47$. Unless otherwise stated, all descriptive statistics and inferential analyses reported in this paper are based on the screened analytical sample ($N = 47$). For analyses involving gender, one participant who selected "prefer not to say" was excluded from subgroup comparisons, yielding 46 cases for gender-based analyses. Sensitivity checks indicated that interpretive conclusions did not materially change when excluded cases were retained; however, reported values reflect the screened dataset to support higher measurement validity.

4. Results

4.1 System Usability Evaluation

4.1.1 Overall usability results

The perceived usability of the integrated badge dashboard (RQ1) was assessed using the SUS. Descriptive statistics are reported in Table 1. The usability score was $M = 60.11$ ($SD = 18.44$), with a median of $Mdn = 62.50$ and scores ranging from 17.50 to 100.00. The distribution approximated normality ($skewness = -0.15, SE = 0.35$; $kurtosis = -0.32, SE = 0.68$), as visualized in Figure 4a.

To contextualize overall usability, scores were classified into commonly used SUS acceptability categories. Specifically, 38.3% ($n = 18$) of participants rated the system as “not acceptable” ($SUS < 51$), 40.4% ($n = 19$) as “marginal” ($SUS 51 - 70$), and 21.3% ($n = 10$) as “acceptable” ($SUS > 70$). Percentile ranks are shown in Figure 4b. The mean falls below the frequently cited benchmark score of 68 and corresponds to an “OK” adjective rating (Bangor, Kortum and Miller, 2009). A one-sample t -test indicated that the observed mean was significantly lower than 68, $t(46) = -2.93, p = .005, Hedges' g = -0.43$.

Table 1: Descriptive Statistics for SUS Scores ($N = 47$)

Measure	Mean	SD	95% CI	Median	Min.	Max.
SUS Score	60.11	18.44	[54.67, 65.55]	62.50	17.50	100.00

4.1.2 Individual item analysis

Item-level descriptive statistics (RQ2) are reported in Table 2, response distributions are shown in Table 3, and aggregated percentage distributions are visualized in Figure 5. The strongest usability dimension was learnability (Q7; $M = 3.57, Mdn = 4.00$), suggesting that the workflow was perceived as relatively easy to learn. Conversely, items implicating overall usage preference (Q1; $M = 3.02$), perceived integration (Q5; $M = 3.09$), and confidence in system usage (Q9; $M = 2.96$) were closer to neutral. Negatively worded items related to technical support requirements (Q4; $M = 1.87, Mdn = 1.00$) and extensive prior learning (Q10; $M = 2.13, Mdn = 2.00$) were rated well below the midpoint, indicating that participants generally did not perceive the system as requiring external assistance or substantial onboarding.

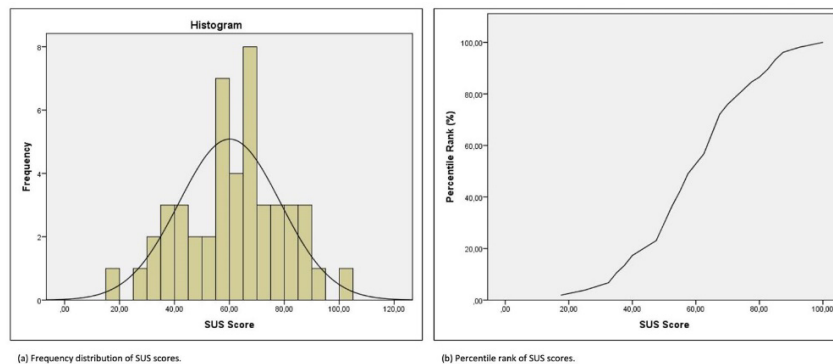


Figure 4: Overview of SUS scores ($N = 47$). The histogram (left) depicts the frequency distribution of SUS scores with a normal density curve. The percentile rank curve (right) presents the cumulative distribution of SUS scores, enabling interpretation of participants' relative standing

Table 2: Descriptive statistics for individual SUS (Brooke, 1996) items. Negative items given in parentheses

Item	Statement	M	SD	Mdn
Q1	I think that I would like to use this badge achievement system frequently.	3.02	1.29	3.00
(Q2)	I found the badge achievement system unnecessarily complex.	2.72	1.25	2.00
Q3	I thought the badge achievement system was easy to use.	3.26	1.26	3.00
(Q4)	I think that I would need the support of a technical person to be able to use this system.	1.87	1.24	1.00
Q5	I found the various functions in this system were well integrated.	3.09	1.10	3.00
(Q6)	I thought there was too much inconsistency in this system.	2.49	1.18	2.00

Q7	would imagine that most people would learn to use this system very quickly.	3.57	1.18	4.00
(Q8)	I found the badge achievement system very cumbersome to use.	2.64	1.19	3.00
Q9	I felt very confident using the badge achievement system.	2.96	1.16	3.00
(Q10)	I needed to learn a lot of things before I could get going with this system.	2.13	1.23	2.00

Table 3: Response distribution for SUS items (%). Negative items given in parentheses.

Item	Strongly Disagree (1)	Disagree (2)	Neutral (3)	Agree (4)	Strongly Agree (5)
Q1	12.8	25.5	25.5	19.1	17.0
(Q2)	12.8	42.6	17.0	14.9	12.8
Q3	12.8	12.8	27.7	29.8	17.0
(Q4)	59.6	12.8	12.8	10.6	4.3
Q5	8.5	21.3	31.9	29.8	8.5
(Q6)	23.4	29.8	27.7	12.8	6.4
Q7	8.5	10.6	14.9	46.8	19.1
(Q8)	19.1	27.7	31.9	12.8	8.5
Q9	12.8	21.3	31.9	25.5	8.5
(Q10)	40.4	29.8	10.6	14.9	4.3

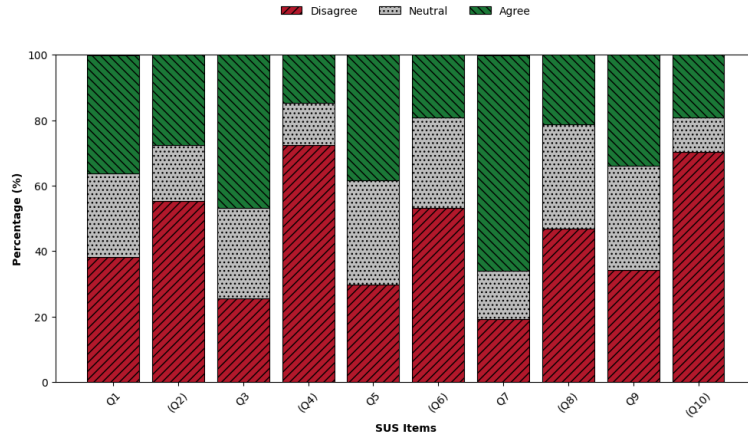


Figure 5: Aggregated response distribution for SUS items. Stacked bars show the percentage of Disagree (strongly disagree + disagree), Neutral, and Agree (agree + strongly agree) responses for each SUS item. Negative items given in parentheses

4.2 Learner Characteristics and Usability

To explore individual differences (*RQ3*), SUS scores were compared by gender, prior badge-system familiarity, and academic performance. Given the exploratory nature of these analyses and the small female subgroup ($n_{female} = 8$), results should be interpreted cautiously. Normality checks using Shapiro-Wilk tests and Q-Q plots indicated no substantial deviations, and parametric comparisons were therefore used with Welch's t test.

Gender-based comparisons were conducted using the 46 participants who reported binary gender categories (38 male, 8 female). Comparisons showed that male participants ($n = 38, M = 61.32, SD = 17.72$) reported slightly higher SUS scores than female participants ($n = 8, M = 59.69, SD = 17.50$), but the difference was not statistically significant, $Welch's t(10.26) = 0.24, p = .816, MD = 1.63, 95\% CI [-13.52, 16.78], Hedges'g = 0.09$.

Participants reporting prior familiarity with other badge systems ($n = 22, M = 63.41, SD = 22.59$) showed higher mean SUS scores than those without prior familiarity ($n = 25, M = 57.20, SD = 13.64$), but this

difference was also not statistically significant, $Welch's t(33.61) = 1.12, p = .270, MD = 6.21, 95\% CI [-5.05, 17.46], Hedges' g = 0.33$.

Academic performance was examined using correlation and simple linear regression analyses ($n = 44$; three participants did not report their GPA). The regression model was not statistically significant, $F(1, 42) = 3.14, p = .083, R^2 = .07$, and GPA was not a significant predictor of SUS scores ($\beta = -.26, t = -1.77, p = .083$).

5. Discussion

This study evaluated the perceived usability of an integrated Blackboard-Notion badge dashboard in an undergraduate engineering course, addressing three research questions concerning overall usability (RQ1), specific usability strengths and weaknesses (RQ2), and the influence of learner characteristics (RQ3). The mean SUS score of 60.11 indicates moderate perceived usability, falling within a marginal acceptability range and significantly below the commonly cited benchmark of 68 (Bangor, Kortum and Miller, 2009). The wide distribution of SUS scores suggests substantial variability in user experiences, consistent with prior findings that gamified systems do not affect all users uniformly (Brambilla et al., 2024). This below-benchmark outcome contrasts with the more favourable usability results reported for single-platform badge implementations (Sousa-Vieira, Ferrero-Castro and López-Ardao, 2022), suggesting that multi-platform architectures may introduce additional interaction costs that are absent in more integrated systems. It is important to interpret these findings as measures of interface usability rather than attitudes toward gamification. The SUS was administered to evaluate the usability of the integrated dashboard used to access badge-related information, rather than the perceived educational effectiveness or motivational value of digital badges, XP points, leaderboards, or the underlying badge design. Consequently, the reported usability outcomes reflect students' experiences interacting with the interface and its integration within the LMS, rather than their opinions regarding the broader badge-based learning approach. This interpretation is consistent with gamification research emphasizing the importance of coherent and meaningful interaction design in sustaining user engagement (Deterding, 2012).

Item-level patterns (RQ2) revealed strengths in learnability (Q7: $M = 3.57$) and a low perceived need for technical or learning support (Q4: $M = 1.87$; Q10: $M = 2.13$). These results suggest that the system effectively supports independent operation, aligning with prior calls for learner-centric control, self-directed progression, and transparent learning pathways within digital badge ecosystems (Gibson et al, 2015; Casilli and Hickey, 2016). Items reflecting perceived complexity (Q2: $M = 2.72$) and inconsistency (Q6: $M = 2.49$) were predominantly rated on the "disagree" end of the scale (55.4% and 53.2%, respectively), indicating that the system was largely viewed as coherent and internally consistent.

However, ratings were more ambivalent for items reflecting *interaction smoothness*, overall preference for using the system (Q1: $M = 3.02$), confidence in using it (Q9: $M = 2.96$), and perceived integration of its components (Q5: $M = 3.09$). This pattern is compatible with a workflow that is learnable yet not fully experienced as seamless, a trade-off that often emerges when LMS platforms embed third-party interfaces with distinct interaction conventions. The neutral integration ratings are particularly noteworthy, given that perceived alignment with existing workflows predicts sustained system use (Hamari, Koivisto and Sarsa, 2014), and engineering students value badges most when they provide professional utility and clear progress feedback, including through structures such as meta-badges (Gregg et al, 2022).

The absence of statistically significant subgroup differences by gender and prior badge familiarity (RQ3) suggests that usability perceptions were not strongly stratified by these characteristics. This finding is consistent with broader gamification research indicating that user responses to gamified educational systems can vary independently of simple demographic characteristics (Hanus and Fox, 2015). One possible explanation is that architectural usability costs affect users relatively uniformly, whereas motivational responses may be more susceptible to individual differences. The weak negative association between GPA and SUS scores warrants cautious interpretation and may reflect differing expectations for system polish, reduced reliance on progress tracking among higher-performing students, or a tendency toward more critical evaluation.

5.1 Design Implications for Multi-platform Badge Dashboards

The results motivate several targeted usability improvements. First, Blackboard should provide stronger signposting and information scent for embedded pages through consistent labels, brief instructional headers, and direct navigation between the catalog and leaderboard. Second, embedded Notion views should be optimized for scanning by minimizing horizontal scroll, prioritizing key columns (earned badges, total XP), and maintaining stable sorting conventions. Third, update semantics should be made more legible through visible

timestamps or explicit refresh mechanisms to increase user confidence, as uncertainty about data currency can reduce trust and perceived usefulness.

More broadly, the findings suggest that the central design question is not whether multi-platform dashboards are usable in principle, but whether the integration provides sufficient *perceived unity* of workflow. Design attention should prioritize reducing friction in recurrent access patterns rather than focusing solely on initial learnability. For institutions, the results highlight that while embedded third-party tools offer rich functionality, they also introduce trade-offs in perceived integration. Institutions must balance these trade-offs against the benefits of enhanced visualization, particularly where seamless LMS integration is a priority.

Finally, the results underscore that usability evaluation should be a standard component of BBL system research. Much of the existing literature focuses on motivational and learning outcomes (Brambilla et al, 2024; Hanus and Fox, 2015) while treating the interface as a neutral delivery mechanism. This study demonstrates that interface characteristics-navigation patterns, perceived integration, interaction confidence are consequential design dimensions that influence whether BBL systems fulfil their intended functions. Future implementations should incorporate usability assessment alongside motivational and performance measures to ensure systems are both pedagogically sound and practically usable.

5.2 Limitations and Future Directions

This study has several limitations. First, the evaluation relied on the SUS, which is a robust global measure but does not identify specific interface breakdowns with the granularity provided by task-based usability tests. Second, the study employed a single-cohort, post-use design without a comparative usability baseline, such as a single-platform implementation, limiting causal inferences about the impact of multi-platform integration. Third, subgroup analyses were exploratory and involved uneven and small group sizes, reducing statistical power to detect differences. Future work should complement SUS findings with task-based measures, observational methods, and usage log data to identify which aspects of the interface contribute to perceived complexity or lower confidence. Comparing LMS-native implementations with embedded third-party dashboards would also help isolate the design decisions that influence integration and usability.

6. Conclusion

This paper presented and evaluated the usability of a pragmatic badge dashboard architecture that embeds Notion database views into Blackboard to present a badge catalog and a leaderboard. The system achieved moderate perceived usability, characterized by comparatively strong learnability and a low perceived need for technical support, but more ambivalent perceptions regarding other aspects of the interaction experience. These findings underscore that usability-centered integration and information architecture are central design requirements for BBL implementations, particularly when they are intended for recurrent use throughout a semester. Overall, the results indicate that multi-platform badge dashboards can achieve acceptable usability in authentic LMS settings, while highlighting opportunities to improve perceived integration and interaction design. The findings reinforce the value of evaluating usability alongside pedagogical outcomes when developing badge-based learning systems.

Acknowledgements

The authors would like to thank their students who participated in this study.

Ethics Declaration: This study was reviewed and approved by the Izmir University of Economics, Faculty of Science and Engineering Research Ethics Committee (06 January 2026; decision no. E-97429853-050.04-114995). Participation was voluntary and based on informed consent. Students were informed that participation or non-participation would not affect their course grades or academic standing and that they could decline participation without penalty. Questionnaire responses were analysed in aggregate, and no personally identifiable information is reported in this paper.

AI Declaration: The authors used generative AI tools to assist in improving the clarity and structure of the manuscript. All content was critically reviewed and fact-checked by the authors to ensure accuracy and originality, and the final manuscript reflects the authors' own interpretations and contributions.

References

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An Empirical Evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction*, 24(6), 574–594. <https://doi.org/10.1080/10447310802205776>

- Bangor, A., Kortum, P. and Miller, J. (2009) 'Determining what individual SUS scores mean: Adding an adjective rating scale', *Journal of Usability Studies*, 4(3), pp. 114–123.
- Brambilla, A., Oliveira, W., Scaico, P. D., Hamari, J., Li, Z., Shi, L. and Dindar, M. (2024) 'The Effects of Gamification on Students' Flow Experience: A Controlled Experimental Study', in *2024 IEEE International Conference on Advanced Learning Technologies (ICALT)*. Los Alamitos, CA: IEEE Computer Society, pp. 245–249. <https://doi.org/10.1109/ICALT61570.2024.00078>
- Brooke, J. (1996) 'SUS: A quick and dirty usability scale', in Jordan, P.W., Thomas, B., Weerdmeester, B.A. and McClelland, I.L. (eds.) *Usability Evaluation in Industry*. London: Taylor & Francis, pp. 189–194.
- Casilli, C. and Hickey, D. (2016) 'Transcending conventional credentialing and assessment paradigms with information-rich digital badges', *The Information Society*, 32(2), pp. 117–129. <https://doi.org/10.1080/01972243.2016.1130500>
- Christy, K.R. and Fox, J. (2014) 'Leaderboards in a Virtual Classroom: A Test of Stereotype Threat and Social Comparison Explanations for Women's Math Performance', *Computers & Education*, 78, pp. 66–77. <https://doi.org/10.1016/j.compedu.2014.05.005>
- Deterding, S. (2012) 'Gamification: Designing for motivation', *Interactions*, 19(4), pp. 14–17. <https://doi.org/10.1145/2212877.2212883>
- 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*. New York, NY: Association for Computing Machinery, pp. 9–15. <https://doi.org/10.1145/2181037.2181040>
- Dicheva, D., Dichev, C., Agre, G. and Angelova, G. (2015) 'Gamification in Education: A Systematic Mapping Study', *Educational Technology & Society*, 18(3), pp. 75–88.
- Elliott, R., Clayton, J. and Iwata, J. (2014) 'Exploring the use of micro-credentialing and digital badges in learning environments to encourage motivation to learn and achieve', *ASCILITE Publications*, pp. 703–707. <https://doi.org/10.14742/apubs.2014.1322>
- Gibson, D., Ostashewski, N., Flintoff, K., Grant, S. and Knight, E. (2015) 'Digital badges in education', *Education and Information Technologies*, 20(2), pp. 403–410. <https://doi.org/10.1007/s10639-013-9291-7>
- Gregg, A., Park, J., Fenton, C., Lang, D. and Handley, M. (2022) 'Exploring the "Why" of Micro-Credentials and Digital Badges: Engineering Students' Motivations for and Perceived Utility of Learning Outside of Class', in *2022 IEEE Frontiers in Education Conference (FIE)*. Piscataway, NJ: IEEE, pp. 1–7. <https://doi.org/10.1109/FIE56618.2022.9962376>
- Hamari, J., Koivisto, J. and Sarsa, H. (2014) 'Does Gamification Work? – A Literature Review of Empirical Studies on Gamification', in *Proceedings of the 47th Hawaii International Conference on System Sciences (HICSS 2014)*. Piscataway, NJ: IEEE, pp. 3025–3034. <https://doi.org/10.1109/HICSS.2014.377>
- 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, pp. 152–161. <https://doi.org/10.1016/j.compedu.2014.08.019>
- Lewis, J.R. and Sauro, J. (2009) 'The Factor Structure of the System Usability Scale', in Kurosu, M. (ed.) *Human Centered Design. HCD 2009*. Lecture Notes in Computer Science, vol. 5619. Berlin, Heidelberg: Springer, pp. 94–103. https://doi.org/10.1007/978-3-642-02806-9_12
- Newby, T.J. and Cheng, Z. (2020) 'Instructional Digital Badges: Effective Learning Tools', *Educational Technology Research and Development*, 68(3), pp. 1053–1067. <https://doi.org/10.1007/s11423-019-09719-7>
- Park, S. and Kim, S. (2019) 'A Badge Design Framework for a Gamified Learning Environment: Cases Analysis and Literature Review for Badge Design', *JMIR Serious Games*, 7(2), e14342. <https://doi.org/10.2196/14342>
- Ryan, R.M. and Deci, E.L. (2000) 'Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being', *American Psychologist*, 55(1), pp. 68–78. <https://doi.org/10.1037/0003-066X.55.1.68>
- Shields, R. and Chugh, R. (2017) 'Digital badges – rewards for learning?', *Education and Information Technologies*, 22(4), pp. 1817–1824. <https://doi.org/10.1007/s10639-016-9521-x>
- Sousa-Vieira, M.-E., Ferrero-Castro, D. and López-Ardao, J.-C. (2022) 'Design, Development and Use of a Digital Badges System in Higher Education', *Applied Sciences*, 12(1), 220. <https://doi.org/10.3390/app12010220>
- Uludağlı, M. Ç. (2025) 'A Preliminary Analysis of Gamification on Assembly Language Programming Education', *European Conference on Games Based Learning*, 19(2), pp. 845–854. <https://doi.org/10.34190/ecgbl.19.2.3886>