

Evaluating Perceived Usability and Gamification Perception in a Lean UX Hotel Booking Prototype

Putri Taqwa Prasetyaningrum¹, Norshahila Ibrahim²

¹Faculty of Information Technology, Universitas Mercu Buana Yogyakarta, Yogyakarta, Indonesia

² Faculty of Computing and Meta-Technology, Sultan Idris Education University, Tanjong Malim, Perak, Malaysia

Received:

May 19, 2026

Revised:

July 25, 2026

Accepted:

September 2, 2026

Published:

October 24, 2026

Corresponding Author:

Author Name*:

Putri Taqwa
Prasetyaningrum

Email*:

putri@mercubuana-
yogya.ac.id

DOI:

10.63158/journalisi.v8i5.1843

© 2026 Journal of
Information Systems and
Informatics. This open
access article is distributed
under a (CC-BY License)



Abstract. This single-group prototype evaluation examines the association between two self-reported post-use perceptions: perceived usability and gamification perception in an early-stage hotel-booking prototype. It does not test whether gamification improves booking performance. A high-fidelity prototype was developed through two Lean User Experience iterations. After the formative first iteration, 150 adults with prior hotel-booking-app experience completed the same ordered scenario remotely through Maze and then answered the System Usability Scale (SUS) and a study-specific ten-item Gamification Perception Questionnaire (GPQ). Objective task-success, completion-time, error, misclick, and abandonment metrics were not retained for analysis. Mean SUS was 89.85 (SD = 11.37; 95% CI [88.02, 91.68]) and the GPQ descriptive mean was 4.67/5 (SD = 0.38; 95% CI [4.61, 4.73]). The two self-reported measures were positively associated ($r = 0.66$, 95% CI [0.56, 0.74], $p < 0.001$). Exploratory gender and age contrasts yielded small effects. The high SUS score represents perceived usability in this sample and does not establish task success or behavioral effectiveness. The contribution is empirical and design-oriented: it documents a Lean UX prototype evaluation in a transactional hospitality context. Interpretation is limited by the single-group design, possible novelty and ceiling effects, self-selection, remote testing, common-method bias, and the status of the GPQ as a study-specific descriptive index that still requires full psychometric validation.

Keywords: Gamification Perception Questionnaire; Hospitality Technology; Lean UX; Prototype Evaluation; Remote Usability Testing; System Usability Scale

1. INTRODUCTION

Digital transformation has changed hotel-booking applications from simple reservation tools into transactional service platforms in which users search by destination and date, compare rooms and policies, make time-sensitive selections, and confirm purchases. Trust, information quality, and low-friction interaction therefore shape booking intention and experience [1], [2]. In this context, perceived usability is important but incomplete: a high questionnaire score does not, by itself, demonstrate successful task completion, efficiency, error prevention, or conversion.

Gamification introduces game-design elements into non-game services to support participation and motivation [7], [8]. In this study, gamification perception is defined narrowly as a user's post-use appraisal of whether the implemented points, badges, challenges, progress cues, leaderboards, and rewards appear appealing, relevant, motivating, and supportive of the prototype experience. It is distinct from behavioral engagement (observable use), enjoyment (an effective response), motivation (a psychological state), and usability (perceived or observed quality of goal achievement). Positive perception therefore cannot be interpreted as evidence that a mechanic improved performance.

Self-Determination Theory (SDT) informed feature selection by linking visible progress and badges to competence, optional challenges and reward choices to autonomy, and social comparison to relatedness [15]-[18]. The Octalysis framework supplied a design vocabulary for development and accomplishment, empowerment, ownership, and social influence [19]. These mappings guided design hypotheses rather than proving psychological effects: progress cues, points, badges, optional challenges, and leaderboards were included because they might support these needs and motivational drivers and then were evaluated as part of the whole prototype.

Lean User Experience (Lean UX) was used because it emphasizes explicit assumptions, rapid prototyping, short feedback cycles, and evidence-informed iteration [20]-[24]. This approach is relevant to booking because motivational features can compete with a user's immediate transactional goal. Unlike banking, where security and financial control dominate, education, where sustained learning is central, or mental-health systems,

where clinical sensitivity and safety are paramount, hotel booking combines inventory comparison, date and price constraints, cancellation policies, and purchase commitment. Gamification must therefore remain subordinate to the booking flow.

Hospitality and tourism studies have examined gamification as a driver of engagement, brand awareness, loyalty, fun, and continued use [26]-[31], but direct evidence on perceived usability during an early hotel-booking prototype evaluation remains limited. The study asks: (RQ1) What level of perceived usability does the final prototype obtain on SUS? (RQ2) How strongly is the study-specific GPQ descriptive index associated with SUS? The contribution is empirical and design-oriented rather than methodological: it documents two Lean UX iterations, reports a single-group post-use evaluation, and derives bounded design priorities without claiming that gamification is effective or causal.

2. RELATED WORK AND THEORETICAL FOUNDATION

2.1. Gamification in digital services

Gamification has been examined in e-commerce, banking, education, health, tourism, and hospitality. E-commerce and banking research links gameful interaction to reputation, repurchase intention, and engagement [11]-[14], whereas hospitality research emphasizes tourism experience, brand loyalty, continued app use, and the meaning of fun [26]-[31]. These outcomes are informative but not interchangeable with the usability of a transactional booking flow. A user may enjoy rewards while still encountering friction in search, room comparison, cancellation review, or confirmation; conversely, an efficient booking flow may produce little interest in social or reward mechanics.

2.2. Usability and the System Usability Scale

Usability concerns effectiveness, efficiency, and satisfaction in a specified context of use [3]. SUS is a widely used ten-item instrument that yields a 0-100 index of perceived usability [4], [5]. It is brief and suitable for prototypes, but its score is not a percentage and must not be equated with task success, completion time, error rate, or accessibility. Normative adjective interpretations can aid communication, although high scores can exhibit ceiling effects and still require behavioral corroboration [6].

2.3. Lean UX as an iterative design approach

Lean UX integrates hypothesis-driven design with rapid experimentation. Teams formulate assumptions, build a minimum viable representation, collect feedback, and revise the design [20]-[24]. In the present work, Iteration 1 was formative and changed navigation, wording, and reward placement; the 150-participant evaluation occurred after Iteration 2. Thus, the study documents evidence-informed design changes but does not estimate an iteration effect because the two versions were not tested in a controlled between-group or repeated-measures comparison.

2.4. Synthesis of prior studies and research gap

Table 1 positions the present work against adjacent digital-service and direct tourism/hospitality studies. Existing evidence supports the relevance of gamification to engagement and loyalty, but most studies do not jointly report Lean UX design changes, perceived usability, a study-specific gamification perception index, and the limitations of an early single-group booking prototype.

Table 1. Positioning of the present study against representative prior work

Study context	Primary focus	Evaluation approach	Gap relevant to this study
E-commerce [11]	Reputation, use, and repurchase intention	Behavioral model and survey	Not a transactional travel prototype
Mobile banking [12]-[14]	Engagement and personalized gamification	Theory-driven survey and Rasch analysis	Different risk, trust, and service context
Tourism/hospitality review [26]	State of digital-platform gamification research	Systematic literature review	Limited task-level usability evidence
Hotel gamified mobile app [27]	Visitor motives and continued use	Qualitative studies	Does not evaluate a booking flow with SUS

Study context	Primary focus	Evaluation approach	Gap relevant to this study
Tourism marketing [28]-[31]	Experience, engagement, brand awareness, and loyalty	Conceptual and survey-based studies	Focuses on marketing outcomes rather than prototype usability
Mental-health support [33]	Exploratory gamified-system design	User-centered exploratory study	Different goals and safety constraints
Present study	Perceived usability and gamification perception	Two Lean UX iterations; single-group SUS-GPQ evaluation	No causal or behavioral-effect claim

2.5. Conceptual relationship and analytical expectations

The conceptual model treats gamification perception and perceived usability as distinct self-reported post-use variables. The GPQ is a study-specific descriptive index of appraisal toward the implemented mechanics; it is not treated as a previously validated measure of engagement, enjoyment, motivation, or usability. SUS captures perceived ease, consistency, learnability, and confidence with the prototype. The analysis therefore asks whether participants who rate the mechanics more positively also tend to report higher perceived usability. A positive GPQ-SUS association is theoretically plausible when progress and feedback are perceived as useful, but the same-session Likert-type measurements create common-method risk [34]. The expectation is therefore associative, not causal. No non-gamified control, baseline condition, or A/B test was included, and demographic analyses are exploratory rather than evidence of universal inclusiveness.

3. METHODS

3.1. Research design

This study used a single-group, cross-sectional evaluation of a high-fidelity prototype. Iteration 1 was a formative research-team cognitive walkthrough and did not contribute a participant-level inferential dataset. All 150 participants evaluated the final Iteration 2

prototype remotely in one post-use session. The design supports estimation of perceived usability and association between self-reported measures, but it cannot establish improvement over Iteration 1 or a causal effect of gamification.

3.2. Lean UX iterations

The development process followed the Think-Design-Test cycle shown in Figure 1. During Think, the team translated booking and reward assumptions into design hypotheses. During Design, a high-fidelity MVP was prepared in Figma. During Test, the research team first conducted a structured cognitive walkthrough of the core paths; the resulting issues informed the second iteration, which was then evaluated by the 150 participants through Maze.

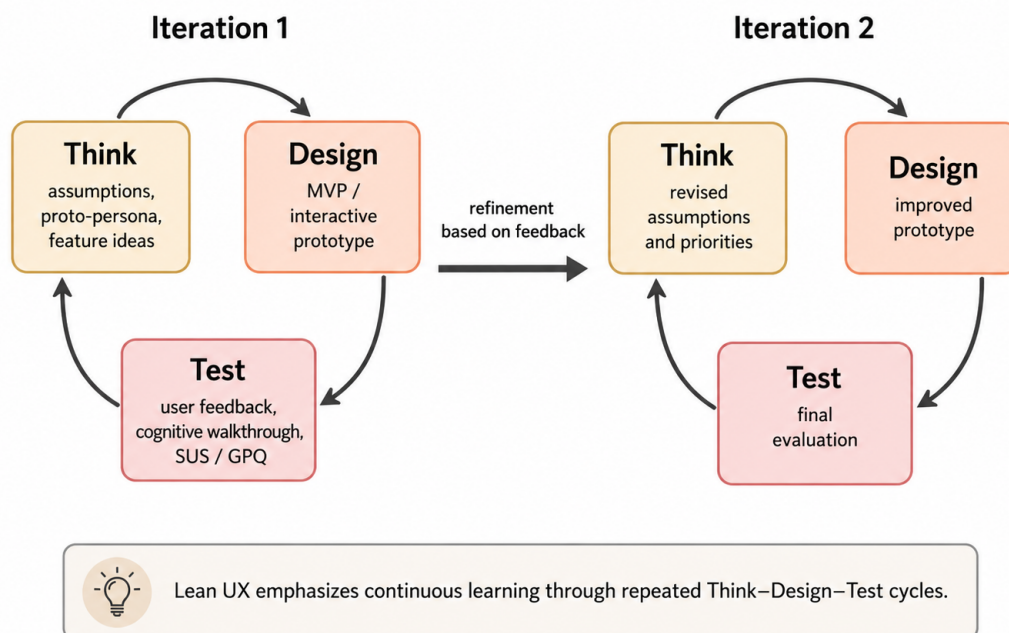


Figure 1. Two-iteration Lean UX cycle; the 150-participant evaluation occurred after Iteration 2.

Iteration 1 focused on sign-in, hotel discovery, room selection, booking review, confirmation, and access to reward information. Walkthrough observations identified navigation ambiguity, dispersed access to rewards, and unclear progress feedback. Iteration 2 clarified labels and next actions, consolidated reward access, and simplified progress cues before the final remote study. Because Iteration 1 was formative and lacked a comparable participant outcome dataset, the manuscript reports design changes rather than statistical improvements between iterations.

Table 2. Summary of the two Lean UX iterations

Iteration	Evaluation focus	Evidence and sample	Design action or analytical role
1	Core booking flow and initial mechanics	Research-team cognitive walkthrough; no participant inferential dataset	Clarified navigation and wording; consolidated reward access; simplified progress cues
2	Revised flow and integrated gamification	Remote final-prototype evaluation; n = 150; SUS, GPQ, and optional comments	Source of all reported quantitative results and the next-iteration backlog

3.3. Prototype and task context

The assigned scenario required each participant, in the same order, to: (T1) access the application and begin a booking; (T2) search for a hotel using destination, date, and guest information; (T3) compare an accommodation option and select a room; (T4) review and confirm the booking flow; and (T5) inspect the points, badges, daily challenge, leaderboard, and reward-redemption areas. Gamification components were positioned around, rather than as mandatory gates within, the booking tasks.

The search-select-review-confirm structure resembles common online travel agency flows, including those used by Traveloka, Booking.com, Agoda, and Airbnb, but the prototype was independently designed and was not a visual replica or a benchmark test of those commercial applications. The proposed visual refinement in Section 5.1 was created after the reported evaluation and was not seen or tested by participants.

3.4. Participants and data collection

A total of 150 participants aged 18-45 years were recruited through online social-media and hotel-booking communities. Eligibility required prior experience with a hotel-booking application. A study-information page preceded testing, and consent confirmation was required before continuing. The final sample included 97 women and 53 men; 80 participants were aged 18-24 years, 60 were aged 25-34 years, and 10 were aged 35 years or older. Country of residence, the specific recruitment channel, frequency of booking-

app use, preferred booking platform, and compensation status were not retained in the study dataset and therefore cannot be reported.

3.5. Walkthrough and remote testing protocol

All participants followed the same ordered scenario through Maze and completed the questionnaires immediately afterward. Maze served as the remote prototype host, but the archived dataset used for this manuscript did not retain task success, completion time, error, misclick, path deviation, or abandonment indicators. The Iteration 1 cognitive-walkthrough notes were produced by the research team by asking whether the goal was evident, the correct action was visible, progress was understandable, and the next state matched expectations. The retained study records do not establish whether Maze was configured to collect these indicators but they were not exported, or whether an earlier export subsequently became unavailable. Their absence is therefore treated as an archival limitation, not as evidence that Maze never generated task analytics. Optional post-use comments were invited, but neither the number of submitted comments nor the complete comment corpus was retained. Consequently, comment frequencies, participant-level linkage, formal themes, and qualitative saturation cannot be reported; only the broad formative issues documented in the design backlog are used to motivate the proposed next iteration.

Table 3. Standardized task sequence and evidence retained

Stage/task	Required participant action	Evidence retained for analysis
Orientation	Read study information and confirm consent	Eligibility and consent confirmation
T1	Access the application and begin a booking	Prototype exposure; no objective task metric retained
T2	Search by destination, date, and guest information	Prototype exposure; no objective task metric retained
T3	Compare an option and select a room	Prototype exposure; no objective task metric retained
T4	Review and confirm the booking flow	Prototype exposure; no objective task metric retained

Stage/task	Required participant action	Evidence retained for analysis
T5	Inspect points, badges, challenge, leaderboard, and rewards	Prototype exposure; no objective task metric retained
Post-use	Complete SUS, GPQ, and optional comments	SUS, GPQ summary, and formative comments

3.6. Instruments

SUS consists of ten statements rated on a five-point scale. For odd-numbered items, 1 is subtracted from the response; for even-numbered items, the response is subtracted from 5. The adjusted scores are summed and multiplied by 2.5 to produce a 0-100 perceived-usability score [4], [5]. Internal consistency was evaluated with Cronbach's alpha. SUS was not used as a substitute for objective effectiveness or efficiency metrics. The ten-item GPQ was developed specifically for this study to describe perceived appeal, motivational relevance, reward/progress support, task focus, and advocacy toward the prototype's game elements. It has not been previously validated. Each item used a five-point agreement scale, and the complete wording is reported in Appendix A. Because the archived analysis file did not include the complete item-response matrix, item-total correlations, GPQ Cronbach's alpha, exploratory factor analysis, and confirmatory factor analysis could not be reproduced. Accordingly, the GPQ is framed as an initial descriptive index rather than a validated psychometric scale.

Table 4. Theory-informed feature hypotheses and GPQ mapping

Design rationale and theory-informed hypothesis	Prototype feature	GPQ items
Competence/development: make progress and achievement visible	Points, progress feedback, and badges	1, 3, 7
Autonomy/empowerment: keep participation optional and choice-oriented	Daily challenges and reward choices	2, 6
Relatedness/social influence: enable comparison without blocking booking	Leaderboard and competition	9
Appeal/meaning: make the reward layer attractive and relevant	Visual presentation and mini-game	4, 5

Design rationale and theory-informed hypothesis	Prototype feature	GPQ items
Task support/advocacy: assess perceived focus and recommendation	Whole gamified experience	8, 10

3.7. Data analysis

Descriptive statistics included means, standard deviations, medians, and benchmark proportions. Two-sided 95% confidence intervals (CIs) for SUS and GPQ means used the *t* distribution; the CI for Pearson's *r* used Fisher's *z* transformation. The GPQ-SUS association was summarized with Pearson's *r* as originally reported. Exploratory gender and age contrasts used Welch's *t*-test because group sizes and variances were unequal, with Cohen's *d* reported as a standardized effect size. The prespecified age contrast combined ages 18-24 and 25-34 (*n* = 140) and compared them with ages 35+ (*n* = 10). Statistical significance was evaluated at $\alpha = 0.05$. No correction was applied because the demographic tests were exploratory. The archived analysis retained the SUS total-score summary and reported internal-consistency coefficient but not the complete SUS item-response table; therefore, SUS item means, item-total correlations, and case-level item profiles could not be reproduced for this revision.

4. RESULTS AND DISCUSSION

4.1. Data screening and analytical scope

All 150 post-use records were retained. The archived manuscript analysis contained total SUS, the reported overall GPQ summary, the reported GPQ-SUS correlation, and demographic summary statistics. It did not contain the complete GPQ item-response matrix, paired observations required to reproduce a scatterplot, task-level Maze analytics, or a formal qualitative codebook. Results are therefore limited to reproducible summary statistics and clearly labeled exploratory comparisons.

4.2. Perceived usability

Mean SUS was 89.85 (*SD* = 11.37; median = 92.50; 95% CI [88.02, 91.68]). In this sample, 97.3% scored at or above 68 and 90.0% at or above 80.3. Cronbach's α was 0.845. The concentration of scores near the upper end indicates a potential ceiling effect; the score is therefore described as high perceived usability rather than proof of high behavioral

effectiveness. Two participants recorded SUS scores of 50 or lower and were retained because no pre-specified exclusion rule applied. This upper-end concentration reduces the instrument's ability to distinguish among participants with already favorable perceptions and may compress associations with other positively rated measures.

Table 5. Descriptive statistics for SUS

Statistic	Value
Mean	89.85
Standard deviation	11.37
95% CI for mean	[88.02, 91.68]
Median	92.50
SUS $\geq$ 68	97.3%
SUS $\geq$ 80.3	90.0%
Cronbach's alpha	0.845

4.3. Gamification perception

The GPQ descriptive mean was 4.67/5 (SD = 0.38; 95% CI [4.61, 4.73]). The three recoverable item means were 4.81 for engagement from points and badges, 4.76 for perceived task focus, and 4.73 for recommendation because of the gamified experience. Table 6 reports all item wordings and explicitly distinguishes available from unavailable item-level statistics. In the absence of the complete item-response matrix, the GPQ is not presented as a psychometrically validated scale.

Table 6. Complete GPQ item wording and recoverable item-level means. NR = not retained in the archived analysis file

Item and study-specific GPQ statement	Mean
1. Points and badges make the application more engaging	4.81
2. Daily challenges motivate me to return to the application	NR
3. The leaderboard helps me understand my progress relative to others	NR
4. The visual presentation of the gamification features is attractive	NR
5. The mini-game makes the hotel-booking experience more enjoyable	NR
6. Collecting badges motivates me to continue using the application	NR
7. Points earned from completing activities make my progress clear	NR

Item and study-specific GPQ statement	Mean
8. Gamification helps me stay focused while using the application	4.76
9. Competition in the application motivates continued participation	NR
10. I would recommend the application because of its gamification features	4.73

The three available item-level results correspond to visible reward points, task-progress feedback, and the integrated reward layer. They indicate positive post-use appraisal but do not isolate the causal effect of any feature, establish long-term engagement, or demonstrate booking performance.

4.4. Association between gamification perception and usability

Pearson's correlation indicated a positive association between the two post-use self-reports, $r = 0.66$, 95% CI [0.56, 0.74], $p < 0.001$, $n = 150$. Participants who rated the gamification layer more positively also tended to report higher perceived usability. Because both measures were collected in the same session using similar Likert-type response scales, common method variance may inflate the association [34]. No causal or feature-level inference is made.

Table 7. Correlation between GPQ and SUS

Association	r	95% CI	p	n
Overall GPQ descriptive index with SUS score	0.66	[0.56, 0.74]	< 0.001	150

4.5. Demographic comparisons

Women reported mean SUS = 90.62 (SD = 10.27), and men reported 88.44 (SD = 13.15). A Welch comparison based on the reported summaries was not significant, $t(87.19) = 1.05$, $p = 0.299$, mean difference = 2.18, 95% CI [-1.97, 6.33], Cohen's $d = 0.19$. For the prespecified age contrast, the combined 18-34 group had mean SUS = 90.05 (SD = 11.62, $n = 140$), compared with 87.60 (SD = 8.91, $n = 10$) for ages 35+; $t(11.31) = 0.82$, $p = 0.428$, mean difference = 2.45, 95% CI [-4.09, 9.00], $d = 0.21$. These small, imprecise effects do not establish inclusiveness, particularly given the small oldest group.

Table 8. SUS scores by demographic group

Group	n	Mean SUS	SD
Women	97	90.62	10.27
Men	53	88.44	13.15
Age 18-24	80	89.81	14.43
Age 25-34	60	90.38	6.26
Age 18-34 combined	140	90.05	11.62
Age >= 35	10	87.60	8.91

Table 9. Exploratory demographic contrasts using Welch's t-test

Contrast	Mean difference	95% CI	t(df)	p	Cohen's d
Women - men	2.18	[-1.97, 6.33]	1.05 (87.19)	0.299	0.19
Age 18-34 - age 35+	2.45	[-4.09, 9.00]	0.82 (11.31)	0.428	0.21

Task-performance outcomes are not reported because the archived Maze output did not retain task success, completion time, error count, misclick rate, path deviation, or abandonment. The two low-SUS cases could not be linked to specific task failures or negative comments. Consequently, the study supports conclusions about self-reported perceptions only and makes no claim about behavioral usability, conversion, or accessibility. Item-level SUS response profiles were likewise unavailable, so localized difficulties hidden by the total score cannot be identified.

4.6. Discussion

The mean SUS score indicates high perceived usability in this sample, but four qualifications are essential. First, participants encountered a high-fidelity prototype rather than an operational booking system. Second, recruitment from online channels among people with prior booking-app experience may have selected digitally confident or gamification-positive respondents. Third, immediate post-use ratings may reflect novelty and demand characteristics. Fourth, the score distribution suggests a ceiling effect, limiting discrimination among highly positive experiences [6]. Remote conditions also prevented consistent observation of context and assistance.

The positive GPQ-SUS association is the principal empirical finding, but it is an association between two self-reported post-use perceptions. Same-session measurement, similar response formats, and a generally positive response tendency create common-method risk [34]. The result is compatible with the SDT-informed design hypothesis that progress and feedback can support competence, yet it cannot demonstrate that points, badges, challenges, or leaderboards led to higher usability. Moreover, the three recoverable GPQ item means cover only a subset of the questionnaire and cannot substitute for complete item-level evidence, reliability analysis, or construct validation of the full instrument.

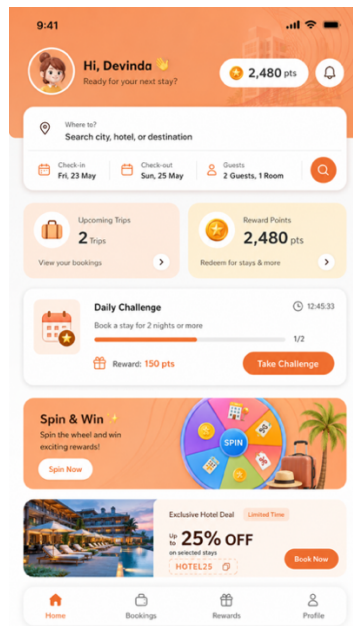
A non-gamified comparison is necessary to determine whether gamification improves, harms, or merely coexists with perceived and behavioral usability. Future A/B testing should keep the booking content and interaction flow constant while varying the gamification layer. Feature-level factorial testing could then distinguish the effects of progress cues, rewards, social comparison, and challenges rather than attributing the overall correlation to a single mechanic.

The next iteration should treat accessibility and potential harm as design requirements. Colorful reward cues require sufficient contrast, non-color status indicators, scalable text, keyboard and screen reader support, reduced-motion options, and a clear focus order consistent with WCAG 2.2 [35]. Gamification can also distract from price and policy review, intensify social-comparison pressure, expose activity through rankings, encourage reward manipulation, or overemphasize points at the expense of informed booking. Optional participation, privacy controls, transparent reward rules, and easy leaderboard opt-out should therefore accompany any further implementation.

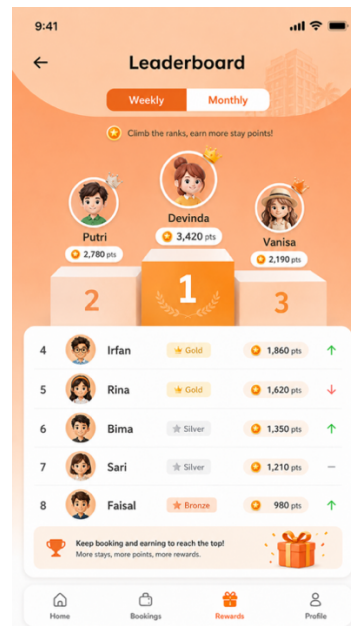
4.6.1. Implications for the next design iteration

The quantitative post-use ratings and formative walkthrough/comment inputs were translated into a prioritized backlog. The comments are not labeled as formal qualitative evidence because no independent coding, intercoder agreement, or saturation procedure was conducted. Figure 2 illustrates a proposed interface refinement derived after the evaluation; it was not included in the tested prototype and is not evidence of improved usability. Because neither the number of comments nor the complete verbatim corpus was retained, no prevalence claim is made about any comment topic.

In Figure 2(a), booking functions remain visually dominant while points, progress, challenges, and promotional rewards are supporting components. Figure 2(b) proposes weekly and monthly ranking, loyalty tiers, and movement indicators. These screens are design hypotheses for a subsequent study. They require a separate accessibility review and controlled evaluation before any performance or usability claim can be made.



(a) home dashboard



(b) Leaderboard

Figure 2. Proposed interface refinement for the next Lean UX iteration: (a) integrated hotel booking and reward dashboard; (b) redesigned leaderboard and loyalty-tier interface.

Table 10. Evidence-based priorities for the next iteration

Priority	Current evidence	Required next evidence
Preserve visible progress and reward feedback	Positive GPQ appraisal; no feature-level effect	A/B test against an otherwise identical non-gamified flow
Keep mechanics outside critical booking steps	High perceived SUS only	Task success, time, errors, misclicks, and abandonment

Priority	Current evidence	Required next evidence
Improve accessibility and user control	No accessibility test was performed	WCAG 2.2 audit, assistive-technology test, opt-out and privacy evaluation
Validate the GPQ	Study-specific descriptive index	Expert content review, item-total analysis, EFA/CFA, reliability, and convergent validity
Measure durable value and harm	One immediate post-use session	Longitudinal return, conversion, redemption, distraction, and social-pressure outcomes

This study has eight material limitations. It evaluated one high-fidelity prototype in a single-group, cross-sectional design; collected SUS and GPQ in the same session; used a study-specific GPQ without complete psychometric validation; lacked a non-gamified comparator; did not retain objective Maze task metrics; did not retain country, recruitment-channel, booking-frequency, preferred-platform, or compensation variables; used an imbalanced age sample with only 10 participants aged 35+; and did not conduct a formal qualitative or accessibility evaluation. The archived analysis file also did not permit reproduction of item-total statistics, EFA, a participant-level scatterplot, or case-level interpretation of the two low SUS observations. Future work should address these limitations with preregistered A/B testing, complete behavioral logs, a validated GPQ, balanced recruitment, longitudinal retention and conversion outcomes, reward-redemption data, and assistive-technology testing.

5. CONCLUSION

This single-group study evaluated perceived usability and gamification perception after remote use of a Lean UX hotel-booking prototype. Mean SUS was 89.85, the study-specific GPQ descriptive mean was 4.67/5, and the two self-reported perceptions were positively associated ($r = 0.66$). The high SUS value indicates perceived usability in this sample only; it does not demonstrate task success, behavioral effectiveness, accessibility, conversion, retention, or a causal benefit of gamification. Before the GPQ can be used as a stable measurement instrument, it requires full psychometric validation, including

expert content review, reliability and item-total analysis, factor-structure assessment, and convergent and discriminant validity. The proposed interface refinement has not been evaluated and remains a design hypothesis. Future work should compare otherwise equivalent gamified and non-gamified versions through preregistered A/B testing; retain complete Maze logs for task success, completion time, errors, misclicks, path deviation, and abandonment; collect longitudinal booking, return, conversion, retention, and reward-redemption outcomes; and conduct accessibility, privacy, manipulation-risk, and potential-harm testing.

ETHICAL CONSIDERATIONS

Participants viewed study information and confirmed voluntary consent before remote testing. No formal institutional ethics approval identifier was available in the study records supplied for this revision. No direct participant identifiers are reported in this article.

DATA AVAILABILITY

Aggregated statistics are reported in the article. Deidentified participant-level data are not publicly deposited; access may be requested from the corresponding author and remains subject to consent, institutional, and data-protection review.

ACKNOWLEDGMENT

The authors thank the participants who contributed to the remote prototype evaluation and the institutions that supported the research collaboration.

REFERENCES

- [1] D. Zhang, P. Wu, and C. Wu, "The role of key online reviews in affecting online hotel booking: An empirical investigation," *Ind. Manag. Data Syst.*, vol. 122, no. 2, pp. 499–520, 2022, doi: 10.1108/IMDS-06-2021-0355.

- [2] P. T. Prasetyaningrum, R. Suryani, and M. Vicky, "Implementation of a web-based photography service booking system to enhance micro-business services," *J. Sustain. Communities Dev.*, vol. 3, no. 1, 2025, doi: 10.15151/SCD.v3i1.24.
- [3] ISO 9241-11:2018, *Ergonomics of Human-System Interaction—Part 11: Usability: Definitions and Concepts*. Geneva, Switzerland: International Organization for Standardization, 2018.
- [4] J. Brooke, "SUS: A quick and dirty usability scale," in *Usability Evaluation in Industry*, P. W. Jordan, B. Thomas, B. A. Weerdmeester, and I. L. McClelland, Eds. London, U.K.: Taylor & Francis, 1996, pp. 189–194.
- [5] J. R. Lewis, "The System Usability Scale: Past, present, and future," *Int. J. Hum.-Comput. Interact.*, vol. 34, no. 7, pp. 577–590, 2018, doi: 10.1080/10447318.2018.1455307.
- [6] A. Bangor, P. T. Kortum, and J. T. Miller, "An empirical evaluation of the System Usability Scale," *Int. J. Hum.-Comput. Interact.*, vol. 24, no. 6, pp. 574–594, 2008, doi: 10.1080/10447310802205776.
- [7] J. Hamari, "Do badges increase user activity? A field experiment on the effects of gamification," *Comput. Hum. Behav.*, vol. 71, pp. 469–478, 2017, doi: 10.1016/j.chb.2015.03.036.
- [8] J. Krath, L. Schürmann, and H. F. O. von Korflesch, "Revealing the theoretical basis of gamification: A systematic review and analysis of theory in research on gamification, serious games and game-based learning," *Comput. Hum. Behav.*, vol. 125, Art. no. 106963, 2021, doi: 10.1016/j.chb.2021.106963.
- [9] S. Trang and W. H. Weiger, "The perils of gamification: Does engaging with gamified services increase users' willingness to disclose personal information?" *Comput. Hum. Behav.*, vol. 116, Art. no. 106644, 2021, doi: 10.1016/j.chb.2020.106644.
- [10] N. J. Thomas and R. Baral, "Mechanism of gamification: Role of flow in the behavioral and emotional pathways of engagement in management education," *Int. J. Manag. Educ.*, vol. 21, no. 1, Art. no. 100718, 2023, doi: 10.1016/j.ijme.2022.100718.
- [11] M. Aparicio, C. J. Costa, and R. Moises, "Gamification and reputation: Key determinants of e-commerce usage and repurchase intention," *Heliyon*, vol. 7, no. 3, Art. no. e06383, 2021, doi: 10.1016/j.heliyon.2021.e06383.
- [12] P. T. Prasetyaningrum, P. Purwanto, and A. F. Rochim, "Enhancing user engagement in mobile banking through personalized gamification: A Cognitive Evaluation Theory approach," *Int. J. Intell. Eng. Syst.*, vol. 17, no. 4, pp. 788–807, 2024, doi: 10.22266/IJIES2024.0831.60.

- [13] P. T. Prasetyaningrum, P. Purwanto, and A. F. Rochim, "Gamification on mobile banking application: A literature review," in *Proc. 1st Int. Conf. Inf. Syst. Inf. Technol. (ICISIT)*, 2022, pp. 222–227, doi: 10.1109/ICISIT54091.2022.9873034.
- [14] P. T. Prasetyaningrum, P. Purwanto, and A. F. Rochim, "Analyzing customer engagement with gamification approach in the banking sector using the Rasch model," *Ingénierie Syst. Inf.*, vol. 29, no. 3, pp. 1051–1064, 2024, doi: 10.18280/isi.290323.
- [15] E. L. Deci and R. M. Ryan, "Cognitive Evaluation Theory," in *Intrinsic Motivation and Self-Determination in Human Behavior*. Boston, MA, USA: Springer, 1985, pp. 43–85, doi: 10.1007/978-1-4899-2271-7_3.
- [16] C. P. Niemiec and R. M. Ryan, "Autonomy, competence, and relatedness in the classroom: Applying self-determination theory to educational practice," *Theory Res. Educ.*, vol. 7, no. 2, pp. 133–144, 2009, doi: 10.1177/1477878509104318.
- [17] R. M. Ryan and E. L. Deci, *Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness*. New York, NY, USA: Guilford Press, 2017.
- [18] E. L. Deci and R. M. Ryan, "The 'what' and 'why' of goal pursuits: Human needs and the self-determination of behavior," *Psychol. Inq.*, vol. 11, no. 4, pp. 227–268, 2000, doi: 10.1207/S15327965PLI1104_01.
- [19] Y.-K. Chou, *Actionable Gamification: Beyond Points, Badges, and Leaderboards*. Milpitas, CA, USA: Octalysis Media, 2015.
- [20] J. Gothelf, *Lean UX: Applying Lean Principles to Improve User Experience*. Sebastopol, CA, USA: O'Reilly Media, 2013.
- [21] J. Gothelf and J. Seiden, *Lean UX: Designing Great Products with Agile Teams*, 3rd ed. Sebastopol, CA, USA: O'Reilly Media, 2021.
- [22] M. M. Alhammad and A. M. Moreno, "Integrating user experience into Agile: An experience report on Lean UX and Scrum," in *Proc. ACM/IEEE 44th Int. Conf. Softw. Eng.: Softw. Eng. Educ. Training*, 2022, pp. 146–157, doi: 10.1145/3510456.3514156.
- [23] M. Isomursu, A. Sirotkin, P. Voltti, and M. Halonen, "User experience design goes Agile in Lean transformation—A case study," in *Proc. Agile Conf.*, 2012, pp. 1–10, doi: 10.1109/Agile.2012.10.
- [24] F. H. Lermen, P. K. de Moura, V. B. Bertoni, P. Graciano, and G. L. Tortorella, "Does maturity level influence the use of Agile UX methods by digital startups? Evaluating design thinking, lean startup, and lean user experience," *Inf. Softw. Technol.*, vol. 154, Art. no. 107107, 2023, doi: 10.1016/j.infsof.2022.107107.

- [25] M. C. Ramos-Vega, V. M. Palma-Morales, D. Pérez-Marín, and J. M. Moguerza, "Stimulating children's engagement with an educational serious videogame using Lean UX co-design," *Entertain. Comput.*, vol. 38, Art. no. 100405, 2021, doi: 10.1016/j.entcom.2021.100405.
- [26] M. G. Pasca, M. F. Renzi, L. Di Pietro, and R. Guglielmetti Mugion, "Gamification in tourism and hospitality research in the era of digital platforms: A systematic literature review," *J. Serv. Theory Pract.*, vol. 31, no. 5, pp. 691–737, 2021, doi: 10.1108/JSTP-05-2020-0094.
- [27] D. Parapanos and E. Michopoulou, "Innovative mobile technology in hotels and the use of gamification," *Tour. Plan. Develop.*, vol. 20, no. 2, pp. 162–187, 2023, doi: 10.1080/21568316.2022.2107563.
- [28] F. Xu, F. Tian, D. Buhalis, J. Weber, and H. Zhang, "Tourists as mobile gamers: Gamification for tourism marketing," *J. Travel Tour. Mark.*, vol. 33, no. 8, pp. 1124–1142, 2016, doi: 10.1080/10548408.2015.1093999.
- [29] F. Xu, D. Buhalis, and J. Weber, "Serious games and the gamification of tourism," *Tour. Manag.*, vol. 60, pp. 244–256, 2017, doi: 10.1016/j.tourman.2016.11.020.
- [30] M. Abou-Shouk and M. Soliman, "The impact of gamification adoption intention on brand awareness and loyalty in tourism: The mediating effect of customer engagement," *J. Destin. Mark. Manag.*, vol. 20, Art. no. 100559, 2021, doi: 10.1016/j.jdmm.2021.100559.
- [31] A. L. Negrusa, V. Toader, A. Sofica, M. F. Tutunea, and R. V. Rus, "Exploring gamification techniques and applications for sustainable tourism," *Sustainability*, vol. 7, no. 8, pp. 11160–11189, 2015, doi: 10.3390/su70811160.
- [32] T. Wolf, W. H. Weiger, and M. Hammerschmidt, "Experiences that matter? The motivational experiences and business outcomes of gamified services," *J. Bus. Res.*, vol. 106, pp. 353–364, 2020, doi: 10.1016/j.jbusres.2018.12.058.
- [33] P. T. Prasetyaningrum, N. Ibrahim, R. Yuniasanti, P. W. Setyaningsih, and I. R. Subagyo, "Designing gamified systems for mental health support: An exploratory study," *J. Inf. Syst. Inform.*, vol. 6, no. 2, pp. 1086–1102, 2024, doi: 10.51519/journalisi.v6i2.760.
- [34] P. M. Podsakoff, S. B. MacKenzie, J.-Y. Lee, and N. P. Podsakoff, "Common method biases in behavioral research: A critical review of the literature and recommended remedies," *J. Appl. Psychol.*, vol. 88, no. 5, pp. 879–903, 2003, doi: 10.1037/0021-9010.88.5.879.

- [35] World Wide Web Consortium, "Web Content Accessibility Guidelines (WCAG) 2.2," *W3C Recommendation*, Oct. 2023. [Online]. Available: <https://www.w3.org/TR/WCAG22/>. Accessed: Aug. 18, 2026.