Jurnal Paedagogy:
Jurnal Penelitian dan Pengembangan Pendidikan
https://e-journal.undikma.ac.id/index.php/pedagogy/index
Email:paedagogy@undikma.ac.id

Vol. 13 No. 3 : July 2026
E-ISSN: 2722-4627
pp. 1032-1046

Game-Based Learning in Biology Education: Designing and Evaluating the BioQuest Mobile Learning Prototype

Nadya Meirizka Putri, Rizki Hikmawan*
Information Systems and Technology Education Study Program,
Universitas Pendidikan Indonesia
*Corresponding Author. Email: hikmariz@upi.edu

Abstract: This study aims to develop and evaluate BioQuest, a mobile educational game prototype for junior high school Biology learning in Indonesia, using a User-Centered Design (UCD) approach. The evaluation was conducted through two iterative cycles focusing on interface design, usability, cognitive load profiles, and the relationship between usability and extraneous cognitive load (ECL). Data were collected using several instruments, including Cognitive Walkthrough (CW), the System Usability Scale (SUS) adapted from task performance on the Maze platform, the User Experience Questionnaire Plus (UEQ+) covering 10 dimensions, a cognitive load questionnaire measuring intrinsic cognitive load (ICL), extraneous cognitive load (ECL), and germane cognitive load (GCL), as well as semi-structured interviews. The first evaluation cycle (UE1) involved 13 participants, while the second cycle (UE2) involved 21 participants, including both returning and newly recruited students. The results showed that UE1 achieved a SUS score of 86, while UE2 achieved a score of 82.5; both scores were categorized as "very good" and exceeded the average usability threshold. All 10 UEQ+ dimensions in both cycles were classified within the very positive category (2.00–2.73 in UE1 and 2.18–2.60 in UE2). The cognitive load profile was optimal across both evaluation cycles, with extraneous cognitive load (ECL) remaining controlled (low in UE1 and moderate in UE2) and germane cognitive load (GCL) consistently high ($M = 5.74$ in UE1; $M = 6.19$ in UE2), indicating strong cognitive engagement during learning activities. A descriptive comparative analysis revealed a negative relationship between usability and ECL across iterations. These findings provide empirical evidence that a UCD-based iterative approach can effectively support the development of mobile educational games with high usability, positive user experience, and a cognitive load profile that promotes meaningful Biology learning at the junior high school level.

Article History
Received: 28-04-2026
Revised: 30-05-2026
Accepted: 19-06-2026
Published: 20-07-2026

Key Words:
Game-Based Learning; Usability; Cognitive Load; User-Centered Design; Biology Learning.

How to Cite: Meirizka Putri, N., & Hikmawan, R. (2026). Game-Based Learning in Biology Education: Designing and Evaluating the BioQuest Mobile Learning Prototype. *Jurnal Paedagogy*, *13*(3), 1032-1046. https://doi.org/10.33394/jp.v13i3.20438

https://doi.org/10.33394/jp.v13i3.20438

Introduction

Digital transformation in education has become a global priority, accelerating significantly since the COVID-19 pandemic (UNESCO, 2023). In Indonesia, the Merdeka Curriculum emphasizes technology-based learning and the development of digital competencies (Kemendikbudristek, 2022). However, Biology learning at the middle school level presents particular challenges, as abstract concepts such as cells, metabolism, and organ systems are difficult to visualize through conventional methods (Zukswert et al., 2020).

Game-Based Learning (GBL) has emerged as an innovative solution by integrating game elements with pedagogical objectives to enhance student motivation, engagement, and learning outcomes (Kalogiannakis et al., 2021). Furthermore, mobile-based GBL platforms extend accessibility, enabling students to learn beyond classroom boundaries while leveraging their familiarity with mobile technology (Papadakis, 2021; Dewi et al., 2021).

Despite its potential, the development of mobile GBL often overlooks cognitive design principles. Curum and Khedo (2021) reported that many mobile learning applications were developed without applying Cognitive Load Theory (CLT), leading to cognitive overload. Toh and Tasir (2024) further demonstrated that cognitive load is a critical issue in mobile learning for Biology students, with extraneous factors from interface design significantly hindering learning performance. Faudzi et al. (2024) further demonstrated that inefficient interface design significantly increases Extraneous Cognitive Load (ECL), which can hinder the learning process.

More critically, Koć-Januchta et al. (2022) identified a strong negative correlation between ECL and usability ($r = -0.76$, $p < 0.01$), indicating that poor interface design directly reduces learning effectiveness. This issue is particularly important for middle school students, who have limited working memory capacity and are more vulnerable to cognitive overload (Toh & Tasir, 2024; Park & Kim, 2024). The User-Centered Design (UCD) approach was selected for this study because it systematically places the user at the center of every design decision through iterative evaluation cycles, enabling the identification and resolution of usability problems based on actual user behavior and cognitive responses. Compared to conventional design methods such as waterfall or expert-centered design, UCD has been demonstrated to be more effective in aligning interface design with users' cognitive characteristics, thereby directly reducing ECL (Teasley et al., 2021; Alsadi & McPhee, 2021).

However, existing studies have not sufficiently examined the relationship between usability and cognitive load particularly ECL in the context of mobile educational games for Biology learning at the middle school level. Empirical studies that simultaneously integrate an iterative UCD process, full cognitive load profiling (ICL, ECL, and GCL), and mobile educational game development for junior high school Biology remain relatively scarce, particularly in the Indonesian context. Therefore, this study aims to develop and evaluate BioQuest, a mobile educational game prototype using a User-Centered Design (UCD) approach, contributing to the literature by providing an integrated, evidence-based framework that connects UCD practice, comprehensive cognitive load profiling, and mobile game-based Biology learning within the Indonesian junior high school context an area that remains largely underexplored in existing research.

Research Method

This study employs a User-Centered Design (UCD) approach to develop and evaluate the BioQuest mobile educational game prototype for junior high school Biology students. The research was carried out through four main stages: (1) preliminary study through literature review and needs analysis; (2) UI/UX design including user personas, learning flowcharts based on Bloom's Taxonomy (Anderson & Krathwohl, 2001), and high-fidelity prototype development using Figma; (3) usability evaluation in two iterative cycles; and (4) data analysis and report preparation.

The prototype integrates gamification elements including a tiered level system (1–10), interactive quizzes, an XP and star reward system, an adventure-style progress map, and a leaderboard. Each screen follows the mobile-first design principle with a clean layout and consistent navigation to minimize ECL (Sweller, 2010). The interface of the initial BioQuest prototype is shown in Figure 2.

Jurnal Paedagogy:
Jurnal Penelitian dan Pengembangan Pendidikan
https://e-journal.undikma.ac.id/index.php/pedagogy/index
Email:paedagogy@undikma.ac.id

Vol. 13 No. 3 : July 2026
E-ISSN: 2722-4627
pp. 1032-1046

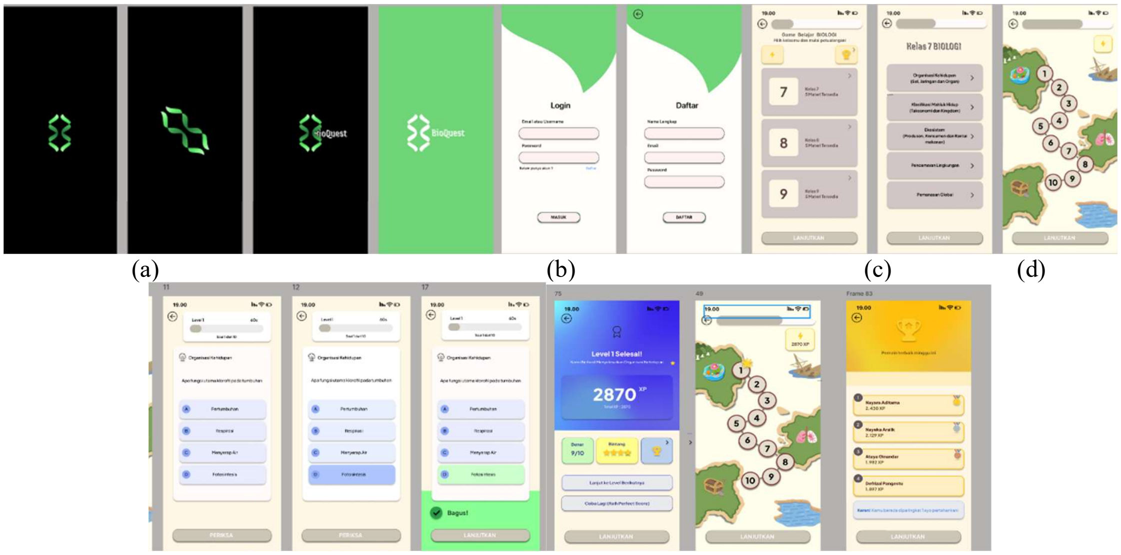


Figure 1. Screenshots of the BioQuest Prototype Interface Before Redesign: (a) Splash Screen and Onboarding, (b) Login/Registration, (c) Class and Subtopic Selection, (d) Level Map, (e) Gameplay and Feedback, (f) Evaluation Results, and (g) Leaderboard.

Usability evaluation was conducted in two iterative cycles, each involving junior high school students selected through purposive sampling based on their familiarity with Android devices and willingness to participate with homeroom teacher consent. Participants in UE2 included both returning participants from UE1 and newly recruited students. Based on UE1 findings, a redesign process guided by CLT principles (Sweller, 2010) and mobile UI/UX design principles (Norman, 2013) was conducted and validated through expert review before UE2.

Instruments

Cognitive Walkthrough

The Cognitive Walkthrough (CW) was used as a performance-based evaluation method to identify usability issues based on task scenarios that represent the actual usage flow of the BioQuest prototype (Hikmawan et al., 2025). The evaluation was conducted by two researchers independently to enhance the reliability of the assessment and minimize observer bias. Participants accessed the Maze platform via a computer or laptop, then interacted directly with the Figma prototype to complete 18 tasks covering all core features of the application, as presented in Table 1. The metrics recorded included task success rate (percentage of tasks successfully completed), drop-off rate (percentage of participants who abandoned a task before completion), misclick rate (number of incorrect clicks per task), and task duration (completion time per task). A think-aloud protocol was implemented so that participants verbalized their thought processes while completing tasks, thereby allowing the identification of cognitive barriers that are not apparent from navigation behavior alone.

Table 1. Task Cognitive Walkthrough Prototype BioQuest

No	Task	Modul/Fitur
1	Log in to the BioQuest app and find the "Organization of Life" material for 7th grade	Authentication
2	Start learning from Level 1 in the "Organization of Life"	Content & Level Selection

Jurnal Paedagogy:
Jurnal Penelitian dan Pengembangan Pendidikan
https://e-journal.undikma.ac.id/index.php/pedagogy/index
Email:paedagogy@undikma.ac.id
Vol. 13 No. 3 : July 2026
E-ISSN: 2722-4627
pp. 1032-1046

	material	
3	For question 1 about the process plants use to make their own food, select the answer "Photosynthesis"	Quiz Module (Multiple Choice)
4	For question 2 about transport tissues, select the answer "Xylem"	Quiz Module (Multiple Choice)
5	After answering question 2 correctly, return to question 2 using the navigation buttons, then select the answer "Epidermis." Observe how the system provides feedback when the answer is incorrect.	Navigation & Feedback Screen
6	For question 3, arrange the following in order from smallest to largest: Cell → Tissue → Organ → Organ System → Organism	Quiz Module (Ordering/Drag Sequence)
7	For question 4, provide an incorrect answer. Notice how the application responds to your error	Error Handling & Feedback System
8	For question 5, read the displayed question. Click the whiteboard icon to view the pre-filled answer.	Essay Question Display
9	For question 6, select the answer "B. An organ is part of an organ system"	Quiz Module (Multiple Choice)
10	For question 7, select the answer "Organization of Life"	Quiz Module (Multiple Choice)
11	For question 8, read the question displayed. Click the board icon to view the pre-filled answer.	Essay Question Display
12	For question 9, try selecting the answer "Wrong." See how the system responds	Error Feedback System
13	For question 10, arrange the numbers in the following order: 3 → 5 → 6 → 4 → 1	Quiz Module (Ordering)
14	After finishing all the questions, find out how you rank compared to other players	Leaderboard Feature
15	Review your results: what score did you get and how many stars did you earn	Score & Achievement Summary
16	Continue to the next level of the Life Organization material	Level Progression System
17	Check the level map. Which levels have you already completed?	Level Map Visualization
18	You want to get a better score. Repeat Level 1 to get a perfect score	Replay Level Feature

Maze Usability Score (MI)

Usability was quantified using the Maze Usability Score (MI), derived from task performance data on the Maze platform. This behavioral approach was selected to capture actual usability performance rather than subjective self-report. It is important to note that this

Jurnal Paedagogy:
Jurnal Penelitian dan Pengembangan Pendidikan
https://e-journal.undikma.ac.id/index.php/pedagogy/index
Email:paedagogy@undikma.ac.id

Vol. 13 No. 3 : July 2026
E-ISSN: 2722-4627
pp. 1032-1046

metric is distinct from the System Usability Scale (SUS) by Brooke (1986), which is a subjective psychometric questionnaire consisting of 10 Likert-scale items. The precise mathematical formula used by the Maze platform to generate the MI score is proprietary and not publicly disclosed; however, the score is derived from task-level performance indicators including success rate, misclick rate, and completion time, and is reported on a 0–100 scale, where higher scores indicate better task-based usability.

User Experience Questionnaire Plus (UEQ+)

The User Experience Questionnaire Plus (UEQ+) by Schrepp et al. (2023) was used to measure the quality of the user experience in a multidimensional manner (RQ2). The UEQ+ is modular, allowing researchers to select the dimensions most relevant to the research context. This study utilized 10 dimensions of the UEQ+ as presented in Table 2, covering pragmatic aspects (Perspicuity, Efficiency, Dependability, Clarity, Comprehensibility) and hedonic aspects (Attractiveness, Stimulation, Novelty, Visual Aesthetics, Usefulness). Participants responded to statements using a 7-point semantic differential scale. Scores for each dimension were transformed to a scale of -3 to +3 and interpreted as: very poor (-3 to -0.8), neutral (-0.8 to 0.8), positive (0.8 to 1.5), or very positive (1.5 to 3).

Table 2. Dimensions and Items of the UEQ+

No	Dimension	Description	Number Items	Example Item
1	Attractiveness	Users' overall impression of the prototype; whether they like it or not	4	Unattractive – Attractive
2	Perspicuity	Ease of learning and understanding how to use the prototype	4	Difficult to understand – Easy to understand
3	Efficiency	Ability to complete tasks quickly and efficiently without excessive effort	4	Slow – Fast
4	Dependability	To what extent users feel they can predict the system's behavior and trust the prototype	4	Unreliable – Reliable
5	Stimulation	To what extent using the prototype feels engaging, enjoyable, and motivating	4	Boring – Engaging– Menarik
6	Novelty	The extent to which the prototype is innovative and creative, thereby capturing users' interest	4	Ordinary – Unique
7	Usefulness	The extent to which the prototype is considered useful and supports users' learning objectives	4	Useless – Useful
8	Clarity	The clarity of information and instructions presented in the prototype interface	4	Unclear – Clear
9	Visual Aesthetics	The overall visual aesthetic quality of the interface	4	Visually unappealing – Visually appealing

Jurnal Paedagogy:
Jurnal Penelitian dan Pengembangan Pendidikan
https://e-journal.undikma.ac.id/index.php/pedagogy/index
Email:paedagogy@undikma.ac.id

Vol. 13 No. 3 : July 2026
E-ISSN: 2722-4627
pp. 1032-1046

10	Comprehensibility	How easily users can understand the content and instructions presented	4	Unintelligible – Intelligible

Cognitive Load Teory (CLT)

The cognitive load questionnaire by Klepsch et al. (2017) was used to measure students' cognitive load profiles while interacting with the BioQuest prototype (RQ3). This instrument consists of 10 items on a 7-point Likert scale (1 = Strongly Disagree; 7 = Strongly Agree), divided into three subscales as presented in Table 3: Intrinsic Cognitive Load (ICL, 3 items), which measures the inherent complexity of the Biology material; Extraneous Cognitive Load (ECL, 4 items), which measures the cognitive load resulting from suboptimal interface design; and Germane Cognitive Load (GCL, 3 items), which measures active cognitive effort in constructing understanding. The average score for each subscale is categorized as low (1–2.9), moderate (3–4.9), or high (5–7). The primary focus of the analysis is on ECL because high ECL scores indicate a suboptimal interface design that needs improvement (Klepsch et al., 2017), whereas high GCL scores are actually expected as they reflect meaningful learning (Sweller, 2010).

Table 3. Cognitive Load Questionnaire Items

No Item	Subscale	Statement	Measurement Focus
1	ICL	The material covered in this game is very complex.	Inherent complexity of the material
2	ICL	The biological concepts presented in this game are difficult to understand.	Difficulty of the material
3	ICL	It takes a lot of effort to understand the material in this game.	Effort required to understand the material
4	ECL	The game's interface design confuses me	Confusion due to UI
5	ECL	The instructions and navigation in this game are hard to follow.	Navigations barriers
6	ECL	I spend a lot of energy figuring out how to use this game, not the material itself.	Interface-related burden
7	ECL	The layout and appearance of this game disrupt my concentration.	Visual/layout distraction
8	GCL	I feel actively engaged in the learning process when using this game.	Active cognitive engagement
9	GCL	I try to connect new material with knowledge I already possess.	Schema construction
10	GCL	Using this game encourages me to think more deeply about the Biology material.	Cognitive elaboratio

Interview

Semi-structured interviews were conducted for 10–15 minutes per participant after all evaluation sessions were completed. The interviews aimed to: (1) explore specific difficulties encountered during the interaction, (2) identify design elements perceived as helpful or

Jurnal Paedagogy:
Jurnal Penelitian dan Pengembangan Pendidikan
https://e-journal.undikma.ac.id/index.php/pedagogy/index
Email:paedagogy@undikma.ac.id

Vol. 13 No. 3 : July 2026
E-ISSN: 2722-4627
pp. 1032-1046

hindering the learning process, (3) explore motivational perceptions of gamification elements, and (4) gather concrete improvement suggestions from the user's perspective. UEQ+ and cognitive load data were collected via Google Forms, which were completed immediately following the cognitive walkthrough session.

Data Analysis

The research data was analyzed descriptively to address RQ1, RQ2, and RQ3. Task performance data from the Cognitive Walkthrough was analyzed based on four key metrics: task success rate, drop-off rate, misclick rate, and task duration. UEQ+ data was analyzed by dimension by calculating average scores and categorizing them based on the UEQ+ interpretation scale and the benchmark database. Cognitive load data was analyzed per sub-scale (ICL, ECL, GCL) and categorized into three levels: low, moderate, or high. Interview data was analyzed thematically to identify patterns of user difficulties and suggestions for improvement, which were then used as the basis for the redesign process in the next iteration. To address RQ4 regarding the relationship between usability and ECL, a comparative descriptive analysis was conducted between the results of UE1 and UE2. This comparison aimed to identify patterns of the relationship between usability levels and ECL scores across iterations, referencing the findings of Koć-Januchta et al. (2022 "21 new participants), who found a strong negative correlation between the two.

Results and Discussion

First User Evaluation

Table 4. Result of User Evaluation 1 (UE1)

						UEQ +		**CLT**		
Task of cw	**Succes rate (%)**	**Drop Off Rate (%)**	**Misclick Rate (%)**	**Duration (seconds)**	**MI**	**Dimension**	**Mean**	**Cognitive Load Dimensions**	**Sesssions 1**	**Category**
1	100	0.0	48.7	51.0	86	Attractiveness	2,29	Intrinsic Cognitive Load (ICL)	2.97	Low
2	100	0.0	63.1	51.7	62	Efficiency	2,33	Germane Cognitive Load (GCL)	5.74	High
3	100	0.0	38.8	19.4	87	Perspicuity	2,25	Extraneous Cognitive Load (ECL)	2.61	Low
4	100	0.0	6.0	10.8	97	Dependability	2,27			
5	100	0.0	77.2	42.6	68	Stimulation	2,60			
6	100	0.0	49.3	24.6	90	Novelty	2,31			
7	100	0.0	45.8	12.6	91	Usefulness	2,73			
8	100	0.0	22.9	12.2	89	Clarity	2,46			
9	100	0.0	48.6	18.3	89	Visual aesthetics	2,29			
10	100	0.0	26.1	8.2	95	Comprehensibility	2,00			
11	100	0.0	6.3	10.1	97					
12	100	0.0	0.0	7.7	10					

Jurnal Paedagogy:
Jurnal Penelitian dan Pengembangan Pendidikan
https://e-journal.undikma.ac.id/index.php/pedagogy/index
Email:paedagogy@undikma.ac.id

Vol. 13 No. 3 : July 2026
E-ISSN: 2722-4627
pp. 1032-1046

					0
13	100	0.0	20.8	20.3	91
14	100	0.0	0.0	6.0	100
15	100	0.0	20.0	12.2	70
16	100	0.0	0.0	6.3	100
17	100	0.0	55.6	7.1	87
18	100	0.0	63.6	32.5	49
Avarage	**100**	**0.0**	**32.9**	**19.6**	**86**

The results of UE1 with 13 participants showed good overall performance—all tasks were completed with a 100% success rate and a SUS score of 86 ("Very Good"), although a misclick rate of 32.9% indicates that the clarity of interactive elements still needs improvement. All 10 UEQ+ dimensions fell within the very positive category (2.00–2.73), with Usefulness and Stimulation as the highest-scoring dimensions reflecting the effectiveness of gamification elements, while Comprehensibility was the lowest-scoring dimension, indicating that the app's instructions still need to be clarified. The cognitive load profile is also considered ideal: both ICL and ECL are low, indicating that the content and interface do not burden users, while the high GCL (M = 5.74) reflects students' active cognitive engagement in the meaningful learning process.

Table 5. UE1 Respondent Suggestions

No	Suggestions
1	Improve the home screen design to make it more visually appealing
2	Add different color themes for each level to indicate progress
3	Add a rating button once students have completed a task
4	Add a "clickable" indicator so students know which elements can be clicked

User feedback obtained from student interviews. The suggestions included improving the home screen's visual clarity, adding color variations for each level, adding a rating button after completing a task, and adding a "clickable" indicator to highlight accessible elements. These suggestions indicate that users still have difficulty recognizing navigation elements and distinguishing interface components. Based on these findings, system improvements are focused on enhancing visual clarity and ease of navigation to make the system more intuitive to use.

UI/UX Redesign

Based on the interview results, the students suggested adding a "click" indicator to interactive elements, improving the home screen layout, adding different color themes for each level, and adding a rating button after completing a task. These suggestions formed the basis for the development and improvement of the BioQuest prototype features, as presented in the following table.

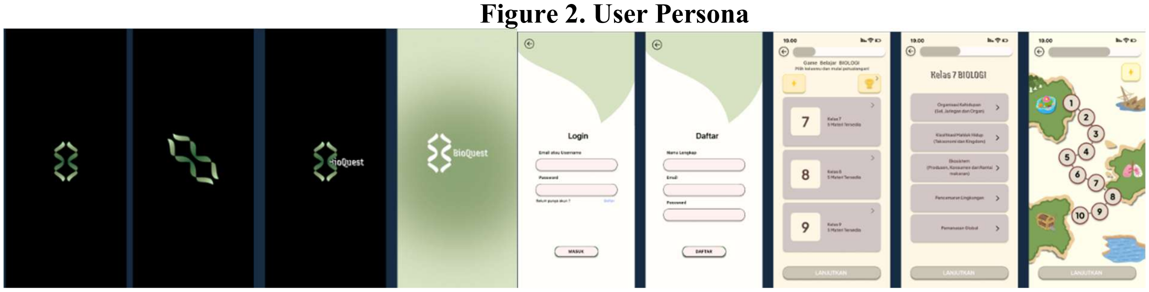
Jurnal Paedagogy:
Jurnal Penelitian dan Pengembangan Pendidikan
https://e-journal.undikma.ac.id/index.php/pedagogy/index
Email:paedagogy@undikma.ac.id

Vol. 13 No. 3 : July 2026
E-ISSN: 2722-4627
pp. 1032-1046

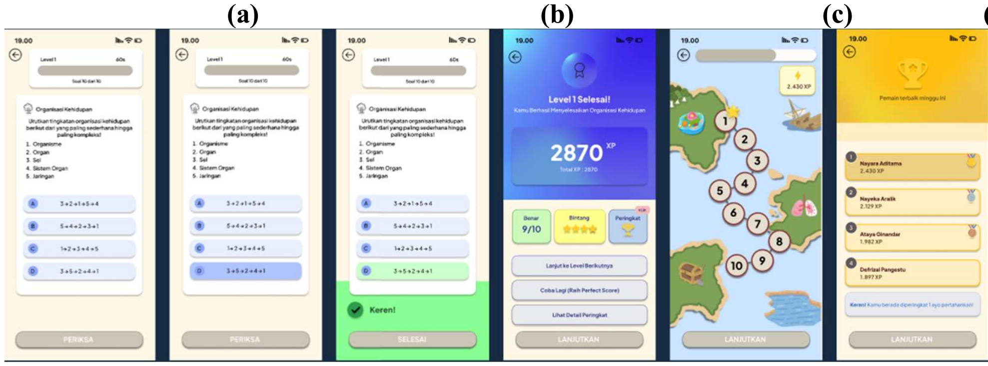
Table 6. Application Features

Feature	Function
Level System (1–10)	Presents interactive questions progressively with locked-level validation
Interactive Quiz	Presents multiple-choice and essay questions with immediate feedback
Reward System (XP & Stars)	Awards XP points and star ratings after completing a level
Progress Map	Displays a visual representation of the learning journey in an adventure-style format
Leaderboard	Displays student rankings based on total XP

Figure 2. User Persona



(a) (b) (c) (d)



(e) (f) (g)

Figure 3. Prototype Interface After Redesign: (a) Splash Screen and Onboarding, (b) Login/Registration, (c) Class and Subtopic Selection, (d) Level Map, (e) Gameplay and Feedback, (f) Evaluation Results, and (g) Leaderboard.

Figure 4 shows the user persona developed as a reference for the BioQuest prototype design and evaluation process. This persona represents the typical characteristics of 9th-grade middle school students the target users including their digital profiles, learning objectives,

Jurnal Paedagogy:
Jurnal Penelitian dan Pengembangan Pendidikan
https://e-journal.undikma.ac.id/index.php/pedagogy/index
Email:paedagogy@undikma.ac.id

Vol. 13 No. 3 : July 2026
E-ISSN: 2722-4627
pp. 1032-1046

and system needs, ensuring that every design decision aligns with the users' cognitive characteristics and real-world needs.

Second User Evaluation

User Evaluation 2 (UE2) was conducted following the redesign of the BioQuest prototype based on findings from UE1. This second evaluation involved 21 participants, comprising both returning participants from UE1 and newly recruited students. Data were collected using instruments identical to those in UE1, including a Cognitive Walkthrough, Maze Usability Score (MI), UEQ+, and a cognitive load questionnaire (CLT).

Table 7. Result of User Evaluation 2 (UE2)

						UEQ +		**CLT**		
Task of cw	**Succes rate (%)**	**Drop Off Rate (%)**	**Misclick Rate (%)**	**Duration (seconds)**	**MI**	**Dimension**	**Mean**	**Cognitive Load Dimensions**	**Sesssions 2**	**Category**
1	100	0.0	11.1	20.4	93	Attractiveness	2,60	Intrinsic Cognitive Load (ICL)	4.22	Moderate
2	94.7	5.3	52.8	25.3	75	Efficiency	2,49	Germane Cognitive Load (GCL)	6.19	High
3	100	0.0	26.4	20.1	82	Perspicuity	2,38	Extraneous Cognitive Load (ECL)	3.23	Moderate
4	100	0.0	5.3	10.1	98	Dependability	2,46			
5	100	0.0	64.1	20.2	78	Stimulation	2,51			
6	100	0.0	6.7	15.4	96	Novelty	2,38			
7	100	0.0	10.6	14.0	98	Usefulness	2,56			
8	100	0.0	18.5	11.8	92	Clarity	2,50			
9	100	0.0	32.3	16.0	88	Visual aesthetics	2,54			
10	100	0.0	7.1	9.7	97	Comprehensibility	2,18			
11	100	0.0	10.5	8.2	95					
12	100	0.0	14.3	6.9	92					
13	100	0.0	9.4	11.0	95					
14	91.7	8.3	41.9	9.0	88					
15	100	0.0	67.6	16.2	59					
16	90.9	9.1	73.2	10.0	85					
17	80.0	20.0	46.3	9.5	38					
18	100	0.0	58.0	40.7	36					
Avarage	**97.6**	**2.37**	**30.8**	**15.2**	**82.5**					

Jurnal Paedagogy:
Jurnal Penelitian dan Pengembangan Pendidikan
https://e-journal.undikma.ac.id/index.php/pedagogy/index
Email:paedagogy@undikma.ac.id

Vol. 13 No. 3 : July 2026
E-ISSN: 2722-4627
pp. 1032-1046

The evaluation results show a consistent improvement from UE1 to UE2. In terms of task performance, completion time decreased from 19.6 seconds to 15.2 seconds, and the click error rate decreased from 32.9% to 30.8%, indicating that the redesign successfully improved navigation efficiency. The Maze Usability Score (MI) decreased slightly from 86 to 82.5, but both remained in the "very good" category and above the standard average (Bangor et al., 2008). All 10 UEQ+ dimensions in both iterations fell into the very positive category. The Attractiveness dimension showed a notable increase from M = 2.29 in UE1 to M = 2.60 in UE2, suggesting that the visual redesign enhanced users' aesthetic perceptions and overall appeal of the application. Comprehensibility remained the lowest dimension despite increasing from M = 2.00 to M = 2.18. In terms of cognitive load, GCL increased from M = 5.74 to M = 6.19, reflecting increasingly strong active cognitive engagement, while ECL rose from low (M = 2.62) to moderate (M = 3.24), likely influenced by differences in participant characteristics and external factors (Sweller, 2010; Klepsch et al., 2017).

There is evidence of a negative correlation between usability and ECL. In UE1, which had higher usability (Maze Usability Score = 86), ECL was recorded as low (M = 2.62), whereas in UE2, where usability was slightly lower (Maze Usability Score = 82.5), ECL increased to the moderate category (M = 3.24), consistent with the findings of Koć-Januchta et al. (2022). However, this analysis is descriptive and comparative in nature, so the relationship between usability and ECL cannot yet be statistically confirmed.

Discussion

This study developed and evaluated the BioQuest mobile educational game prototype through two iteration cycles based on a User-Centered Design (UCD) approach. Overall, the evaluation results indicate that the developed prototype successfully met adequate usability and user experience standards, while also supporting students' active cognitive engagement in biology learning.

In terms of interface design (RQ1), the BioQuest prototype integrates gamification elements such as a tiered level system, interactive quizzes, an XP and star reward system, an adventure-style progress map, and a leaderboard. This approach aligns with the principles of game-based learning, which emphasize the importance of game elements in fostering student motivation and engagement (Kalogiannakis et al., 2021). Zourmpakis et al. (2023) confirm that adaptive gamification elements, such as reward systems and progress tracking, have been shown to enhance students' intrinsic motivation in science learning by creating a feedback loop that reinforces learning behavior sustainably. An interface design that applies the mobile-first principle with a clean layout and consistent navigation also contributes to reducing ECL, in line with the recommendation by Faudzi et al. (2024) that an uncluttered layout and intuitive navigation are key factors in minimizing external cognitive load in mobile learning applications.

In terms of usability (RQ2), the Maze Usability Score (MI) scores for both iterations 86 (UE1) and 82.5 (UE2) fell into the "Very Good" category and exceeded the average threshold (Bangor et al., 2008). These scores are consistent with the findings of Alsadi and McPhee (2021), which indicate that systems developed through a UCD approach tend to yield high usability scores as the iterative design process allows for the identification and resolution of usability issues prior to full implementation. A slight decrease in the Maze Usability Score (MI) score from UE1 to UE2 can be explained by differences in participant composition—UE2 involved new students who had never interacted with the prototype, so their initial learning curve also influenced overall usability perceptions. Nevertheless, the

Vol. 13 No. 3 : July 2026
E-ISSN: 2722-4627
pp. 1032-1046

consistency of above-average scores across iterations indicates that UCD-based redesign successfully maintained the prototype's usability quality.

In terms of user experience, all 10 UEQ+ dimensions in both iterations fell into the "very positive" category. The improvement in the Attractiveness dimension from UE1 to UE2 indicates that visual enhancements during the redesign process successfully increased the interface's appeal this can be explained by the aesthetic-usability effect theory, where visually more appealing interfaces tend to be perceived as easier to use and trigger positive affect that supports learning motivation (Barz et al., 2024). Conversely, the Comprehensibility dimension, which consistently ranked lowest in both iterations, indicates that the clarity of instructions remains an unresolved design challenge. This phenomenon aligns with the argument by Clark and Mayer (2019) that non-explicit instructions force users to allocate working memory capacity to interpret navigation cues, rather than processing learning content—a condition that directly hinders learning effectiveness.

The cognitive load profile (RQ3) shows a pattern that supports the effectiveness of the prototype design. The increase in ICL from the low to the moderate category indicates that UE2 participants perceived the Biology content as more challenging, which can be scientifically attributed to differences in the participants' academic backgrounds and prior knowledge levels—consistent with the CLT principle that ICL is determined by the inherent complexity of the material relative to the user's prior knowledge (Sweller, 2010). Most significantly, the increase in GCL from $M = 5.74$ (UE1) to $M = 6.19$ (UE2)—both in the high category—reflects the growing strength of students' active cognitive engagement in the knowledge-construction process. This mechanism occurs because gamification elements such as level systems and immediate feedback encourage students to continuously process and consolidate new information with existing schemas (Ghoulam et al., 2024), in line with Mayer (2019) who states that multimedia design following cognitive principles can maximize germane load while minimizing extraneous load.

Regarding the relationship between usability and ECL (RQ4), comparisons across iterations indicate a negative pattern—when usability decreased slightly (SUS from 86 to 82.5), ECL tended to increase (from $M = 2.62$ to $M = 3.24$). This pattern is consistent with the findings of Koć-Januchta et al. (2022), who found a strong negative correlation between usability and ECL ($r = -0.76$, $p < 0.01$), as well as the argument by Hollender et al. (2010) that educational software with high usability inherently results in lower ECL. Scientifically, the increase in ECL on UE2 can be explained by differences in participant composition: the presence of new students who had never interacted with the prototype caused part of their working memory capacity to be allocated to understanding the interface, rather than processing the learning content. This aligns with the CLT principle that prior knowledge plays a crucial role in reducing cognitive load—learners with lower prior knowledge must process more elements simultaneously, thereby increasing cognitive load even if the interface design remains unchanged. It is important to note that although ECL increased, its value remained in the moderate category and did not reach the high category—given that poor UI design can significantly increase ECL to the point of hindering the learning process, these results actually indicate that the interface redesign successfully prevented cognitive overload, even among participants with more diverse backgrounds.

Conclusion

This study successfully addressed the four research questions formulated through the development and evaluation of two iterative cycles of the BioQuest mobile educational game prototype, based on a User-Centered Design approach. In terms of interface design (RQ1), the integration of gamification elements including a tiered level system, interactive quizzes, an XP and star reward system, an adventure-style progress map, and a leaderboard — was proven to create a motivating learning environment and reduce ECL through the application of clean mobile-first design principles and consistent navigation. From the usability perspective (RQ2), the BioQuest prototype achieved a Maze Usability Score (MI) score in the "Very Good" category in both evaluation cycles (UE1 = 86; UE2 = 82.5), exceeding the average threshold, and maintaining all 10 UEQ+ dimensions in the highly positive category, confirming that the UCD-based redesign process successfully maintained consistent user experience quality. The cognitive load profile (RQ3) indicates conditions conducive to meaningful learning: ECL is controlled within low-to-moderate limits, while GCL increased from M = 5.74 (UE1) to M = 6.19 (UE2) both in the high category indicating a strengthening of students' active cognitive engagement in the knowledge-construction process. Regarding the relationship between usability and ECL (RQ4), comparisons across iterations indicate a consistent negative pattern: a decrease in usability correlates with an increase in ECL, although the rise in ECL in UE2 is also attributable to the greater heterogeneity of participants' prior knowledge. Overall, this study provides empirical evidence that a UCD approach with an iterative evaluation cycle is capable of producing mobile educational games with high usability, a positive user experience, and an optimal cognitive load profile to support Biology learning at the junior high school level.

Recommendation

Based on the research findings, recommendations are proposed in two areas. In terms of practical and product development, the full development of BioQuest should prioritize improving the Comprehensibility dimension by adding interactive onboarding tutorials and more explicit navigation signage. The level map and leaderboard should also be refined with clearer progress indicators and differentiated color themes per level. Additionally, a rating or feedback button after each level completion should be incorporated to increase user engagement and support ongoing iterative improvements.

Regarding recommendations for future research, usability and ECL data should be collected at the individual level to enable statistical testing of their relationship using Pearson or Spearman correlation analyses. Evaluations should also be conducted during full implementation with a larger, more homogeneous sample in terms of prior knowledge to improve generalizability. Furthermore, incorporating a control group design would better isolate the effect of the BioQuest prototype on cognitive load and learning outcomes from individual learner differences.

References

Barić, Z. (2024). Exploring the Teaching of Solfeggio, Harmony and Counterpoint in Croatian Music Schools. *Croatian Journal of Education*, *26*(1), 233–280. https://doi.org/10.15516/cje.v26i1.4975

Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives* (Complete ed.). Longman.

Vol. 13 No. 3 : July 2026
E-ISSN: 2722-4627
pp. 1032-1046

Alsadi, M., & McPhee, C. (2021). A user-centered design approach for developing an e-learning system for use in a Saudi Arabian university. *International Journal of Emerging Technologies in Learning (iJET)*, *16*(6), 136–155. https://doi.org/10.3991/ijet.v16i06.16944

Bangor, A., Kortum, P., & Miller, J. (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

Barz, N., Benick, M., Dörrenbächer-Ulrich, L., & Perels, F. (2024). The effect of digital game-based learning interventions on cognitive, metacognitive, and affective-motivational learning outcomes in school: A meta-analysis. *Review of Educational Research*, *94*(2), 193–227. https://doi.org/10.3102/00346543231167795

Clark, R. C., & Mayer, R. E. (2019). *E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning* (4th ed.). Wiley.

Curum, B., & Khedo, K. K. (2021). Cognitive load management in mobile learning systems: A systematic review of empirical studies. *Journal of Computers in Education*, *8*(2), 239–281. https://doi.org/10.1007/s40692-020-00181-6

Dewi, C., Prabawa, H. W., & Suhandi, A. (2021). Mobile-based game-based learning for science: Effectiveness and students' perceptions. *Journal of Physics: Conference Series*, *1806*(1), 012106. https://doi.org/10.1088/1742-6596/1806/1/012106

Faudzi, M. A. Z., Che Cob, Z., Sharudin, S. A., Omar, R., & Ghazali, M. (2024). The effects of user interface design for mobile learning application on learner's extraneous cognitive load: A conceptual framework. In *Proceedings of the Asian HCI Symposium 2023 (Asian CHI '23)* (pp. 51–57). Association for Computing Machinery. https://doi.org/10.1145/3604571.3604579 .

Ghoulam, K., Bouikhalene, B., Babori, A., & Falih, N. (2024). Gamification in e-learning: Bridging educational gaps in developing countries. *International Journal of Advanced Corporate Learning (iJAC)*, *17*(1), 85–95. https://doi.org/10.3991/ijac.v17i1.47631

Hollender, N., Hofmann, C., Deneke, M., & Schmitz, B. (2010). Integrating cognitive load theory and concepts of human–computer interaction. *Computers in Human Behavior*, *26*(6), 1278–1288. https://doi.org/10.1016/j.chb.2010.05.031

Hikmawan, R., Rohendi, D., Septiadi, J., & Arrumaisha, N. (2025). Designing AI-based learning assistant to elevate computational thinking: A UX-focused redesign approach. *Jurnal Paedagogy*, *12*(4), 993–1006. https://doi.org/10.33394/jp.v12i4.16705

Kalogiannakis, M., Papadakis, S., & Zourmpakis, A.-I. (2021). Gamification in science education: A systematic review of the literature. *Education Sciences*, *11*(1), Article 22. https://doi.org/10.3390/educsci11010022

Kemendikbudristek. (2022). *Keputusan Menteri Pendidikan, Kebudayaan, Riset, dan Teknologi Nomor 56/M/2022 tentang Pedoman Penerapan Kurikulum dalam rangka Pemulihan Pembelajaran*. Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia.

Klepsch, M., Schmitz, F., & Seufert, T. (2017). Development and validation of two instruments measuring intrinsic, extraneous, and germane cognitive load. *Frontiers in Psychology*, *8*, Article 1997. https://doi.org/10.3389/fpsyg.2017.01997

Koć-Januchta, M. M., Schönborn, K. J., Roehrig, C., Tibell, L. A. E., & Heller, C. (2022). "Connecting concepts helps put main ideas together": Cognitive load and usability in learning biology with an AI-enriched textbook. *International Journal of Educational*

Vol. 13 No. 3 : July 2026
E-ISSN: 2722-4627
pp. 1032-1046

Technology in Higher Education, *19*, Article 11. https://doi.org/10.1186/s41239-021-00317-3

Mayer, R. E. (2019). *Multimedia learning* (3rd ed.). Cambridge University Press.

Norman, D. (2013). *The design of everyday things* (Revised and expanded ed.). Basic Books.

Papadakis, S. (2021). Tools for evaluating educational apps for young children: A systematic review of the literature. *Advances in Mobile Learning Educational Research*, *1*(1), 1–10. https://doi.org/10.25082/AMLER.2021.01.001

Park, H., & Kim, S.-Y. (2024). Differential impact of WM load on attention in young adults versus children and adolescents. *Children*, *11*(9), Article 1057. https://doi.org/10.3390/children11091057

Schrepp, M., & Thomaschewski, J. (2023). *UEQ+ handbook* (Version 6). https://ueqplus.ueq-research.org/Material/UEQ+_Handbook_V6.pdf

Schrepp, M., Kollmorgen, J., Meiners, A.-L., Hinderks, A., Winter, D., Santoso, H. B., & Thomaschewski, J. (2023). On the importance of UX quality aspects for different product categories. *International Journal of Interactive Multimedia and Artificial Intelligence*, *8*(3), 50–57. https://doi.org/10.9781/ijimai.2023.03.001

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review*, *22*(2), 123–138. https://doi.org/10.1007/s10648-010-9128-5

Teasley, S. D., Kay, M., Elkins, S., & Hammond, J. (2021). User-centered design for a student-facing dashboard grounded in learning theory. In M. Sahin & D. Ifenthaler (Eds.), *Visualizations and dashboards for learning analytics* (pp. 151–175). Springer. https://doi.org/10.1007/978-3-030-81222-5_9

Toh, T. J., & Tasir, Z. (2024). The impact of a mobile learning application on students' cognitive load and learning performance in biology. *Journal of Information Technology Education: Research*, *23*, Article 26. https://doi.org/10.28945/5380

UNESCO. (2023). *Technology in education: A tool on whose terms? Global education monitoring report 2023*. UNESCO. https://doi.org/10.54675/UZQV8501

Zourmpakis, A.-I., Papadakis, S., & Kalogiannakis, M. (2023). Adaptive gamification in science education: An analysis of the impact of implementation and adapted game elements on students' motivation. *Computers*, *12*(7), Article 143. https://doi.org/10.3390/computers12070143

Zukswert, J. M., Barton, H. A., & Momsen, J. L. (2020). Troublesome jargon in biology: What terms do students find difficult? *CBE—Life Sciences Education*, *19*(1), Article ar6. https://doi.org/10.1187/cbe.19-04-0093