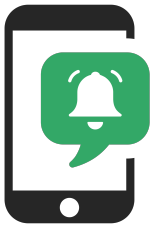

DIGI-Teens

WP3 Evaluation

D3.1.1 Formative Evaluation of the Proof of Concept

https://digiteens.polito.it

Finanziato dall'Unione europea
NextGenerationEU

This project has received funding from the European Union – Next Generation EU Mission 4 Component 1, within the PRIN 2022 program (CUP E53D23007840006).

CUP No.: E53D23007840006
Deliverable: D3.1.1 Formative Evaluation of the Proof of Concept

Project Start Date: 28/09/2023 **Duration**: 29 months
Coordinator: *Politecnico di Torino, Italy*

Deliverable No:	D3.1.1
WP No:	3
WP Leader:	Catia Prandi
Final version (date):	01/03/2026

Dissemination Level:

PU	Public Use	X
PP	Restricted to other programme participants	
RE	Restricted to a group specified by the consortium	
CO	Confidential, only for members of the consortium	

DOCUMENT SUMMARY INFORMATION

Project title:	**Improving Digital Wellbeing With and For Teens: a Gamified and Personalized Intelligent System**
Short project name:	DIGI-Teens
CUP No.:	E53D23007840006
Program:	PRIN 2022
Start date of the project:	28/09/2023
End date of the project:	28/02/2026
Project website:	https://digiteens.polito.it

D3.1.1 Formative Evaluation of the Proof of Concept

Work Package:	WP3 Evaluation
Deliverable number:	D3.1.1
Deliverable title:	Formative Evaluation of the Proof of Concept
Final version (date):	01/03/2026
Authors:	Luca Scibetta, Francesco Ballarini, Alberto Monge Roffarello, Catia Prandi, Luigi De Russis
Dissemination Level:	PU
No. pages:	20
Delivery date:	18/09/2025
Responsible person:	L. De Russis
Status:	Final

COPYRIGHT

ACKNOWLEDGEMENTS

“Improving Digital Wellbeing With and For Teens: a Gamified and Personalized Intelligent System” (DIGI-Teens) is a project that has received funding from the European Union – Next Generation EU Mission 4 Component 1, within the PRIN 2022 program (CUP E53D23007840006).

Please see https://digiteens.polito.it for more information.

The partners in the project are Politecnico di Torino and Alma Mater Studiorum - Università di Bologna. The content of this document is the result of the work developed by the partners in the context of the project.

DISCLAIMER

Contents

1 Introduction and Objectives **6**

2 Methodology **6**
2.1 Evaluation Instruments 7
2.1.1 Student Questionnaire 7
2.1.2 Teacher Questionnaire 9

3 Participants **11**

4 Results **12**
4.0.1 System Usability Scale (SUS) 12
4.0.2 User Experience Questionnaire – Short Version (UEQ-S) 13
4.0.3 App-Specific Usability Items 13
4.0.4 Qualitative Feedback 16

5 Conclusion **17**

1 Introduction and Objectives

This deliverable presents the first evaluation conducted on the initial version of the *Proof of Concept* (POC) developed within the project. The evaluation was carried out prior to the final iteration of the platform and served as a foundational step in the iterative refinement process, providing empirical evidence that directly informed the subsequent development cycle. Rather than an in-the-wild deployment, the evaluation adopted a controlled, in-the-lab methodology, structured across two distinct testing sessions in order to systematically assess the platform's core functionalities and usability under supervised conditions.

The evaluation involved young adult participants assuming the roles of the platform's intended end users, namely **students** and **teachers**, ensuring that feedback was collected from perspectives representative of both primary user groups. A **mixed-methods approach** was employed, combining quantitative and qualitative feedback in the **post-questionnaire**, so as to capture both measurable usability indicators and richer, more nuanced accounts of user experience. Observational data gathered during the sessions complemented the participants' self-reported responses, offering additional insight into interaction patterns and potential areas of friction.

The findings emerging from this evaluation were systematically analysed and translated into actionable design and development recommendations. In this way, the evaluation fulfilled a twofold purpose: to assess the alignment of the first version of the platform with the project's intended goals, and to **support the iterative improvement** of the POC, bridging the gap between the initial prototype and the more refined version presented in subsequent deliverables.

2 Methodology

The first evaluation of the platform was conducted in a controlled, in-the-lab setting, involving two independent groups of participants across three structured testing sessions each. The primary objective of this evaluation was to assess the usability, user experience, and overall clarity of both the student and teacher interfaces of the platform in a realistic, simulated classroom scenario, and to gather actionable feedback to guide the subsequent development iteration.

Each session followed a consistent protocol. Within each group, one participant assumed the role of **teacher**, interacting with the dedicated web application to plan and manage classroom activities, while the remaining participants acted as **students**, engaging with the platform from the learner's perspective. At the beginning of each session, participants were given a brief familiarization period to explore their newly assigned accounts before proceeding with the tasks. Each session comprised **three activities**, which were planned by the teacher as indicated by the facilitators and subsequently carried out by the student participants.

Throughout the sessions, participants were invited to actively take notes on any aspects of the platform they found noteworthy, whether positive or negative, to enrich the subsequent feedback collection phase. At the conclusion of all three sessions, every participant was asked to

complete a **questionnaire focused on the student application**. Those who had taken on the role of teacher in at least one session were additionally required to complete a **separate questionnaire addressing the teacher application**. This dual-questionnaire structure allowed for the collection of differentiated feedback reflecting the distinct roles and interaction modalities experienced by participants.

The evaluation thus adopted a **mixed-methods approach**, combining quantitative ratings derived from validated usability and user experience instruments with qualitative open-ended responses and observational data, in order to produce a comprehensive picture of the platform's strengths and areas for improvement.

2.1 Evaluation Instruments

The feedback collected at the end of the evaluation sessions was structured into two separate questionnaires: one targeting the **student application**, administered to all participants, and one targeting the **teacher application**, administered exclusively to those who had performed the teacher role. Both questionnaires were designed by integrating two widely adopted and scientifically validated instruments:

- **System Usability Scale** (SUS), consisting of 10 items rated on a 5-point Likert scale (1 = Strongly disagree, 5 = Strongly agree), used to assess perceived usability across dimensions such as ease of use, complexity, integration, and learnability [3].
- **User Experience Questionnaire – Short Version** (UEQ-S), consisting of 8 bipolar adjective pairs rated on a 7-point semantic differential scale, designed to measure users' perceptions of both pragmatic quality (Obstructive–Supportive, Complicated–Easy, Inefficient–Efficient, Confusing–Clear) and hedonic quality (Boring–Exciting, Uninteresting–Interesting, Conventional–Original, Usual–Modern) [2].

These instruments were complemented by application-specific questions and open-ended sections for qualitative feedback. This combination was chosen for its capacity to produce a large volume of quantifiable data while retaining sensitivity to usability and engagement dimensions, particularly relevant given the student-facing nature of the application.

2.1.1 Student Questionnaire

After completing the testing sessions, all participants were asked to fill in a questionnaire aimed at evaluating their experience with the student application. The questionnaire was organized into five sections, combining validated psychometric instruments with application-specific items and open-ended feedback fields.

Beyond the standardized measures, the questionnaire also included a set of application-specific items rated on a 5-point scale (1 = Not clear at all, 5 = Very clear), aimed at capturing participants' subjective perception of key functional areas of the student application: the overall ease of completing activities, the clarity of the activity flow (planned, ongoing, and completed), the intelligibility of the gamification elements (badges, coins, shop, and profile customization), and the clarity of the Results panel. A multiple-choice item additionally allowed participants to flag specific activities whose results they found unclear, with an optional open-ended follow-up field. Finally, participants were given the opportunity to report

any technical issues encountered during the sessions and to provide open-ended feedback through optional free-text fields.

The full list of questionnaire items is reported in Table 1.

Table 1: Student questionnaire items

Section	Items
Demographic Information	1. Age range 2. Gender 3. Occupation 4. Background 5. Which of the three sessions did you attend as a student? (multiple choice)
System Usability Scale (SUS)	1. I think that I would like to use this system frequently. 2. I found the system unnecessarily complex. 3. I thought the system was easy to use. 4. I think that I would need the support of a technical person to be able to use this system. 5. I found the various functions in this system were well integrated. 6. I thought there was too much inconsistency in this system. 7. I would imagine that most people would learn to use this system very quickly. 8. I found the system very cumbersome to use. 9. I felt very confident using the system. 10. I needed to learn a lot of things before I could get going with this system.
User Experience Questionnaire – Short Version (UEQ-S)	1. Obstructive – Supportive 2. Complicated – Easy 3. Inefficient – Efficient 4. Confusing – Clear 5. Boring – Exciting 6. Not interesting – Interesting 7. Conventional – Original 8. Usual – Modern

Section	Items
App-Specific Usability Items	1. How easy was it to use the app to complete the activities? 2. How clear was the full flow of viewing planned activities, ongoing activities, completing them, and viewing results? 3. How clear was the functioning of the gamification elements (badges, coins, shop, profile customisation)? 4. How clear did you find the Results panel, considering all activities? 5. Was there any activity whose results were particularly unclear? (multiple choice) 6. If so, what issues did you encounter? (open-ended, optional)
Technical Issues	1. Did you encounter any technical bugs during use? (open-ended, optional) 2. Did you experience any stability or performance issues? (open-ended, optional)
Open-Ended Feedback	1. What struck you most about this system? (optional) 2. What would you change in the student app? (optional) 3. Any other comments or general suggestions? (optional)

2.1.2 Teacher Questionnaire

Participants who had taken on the role of teacher in at least one of the sessions were additionally asked to complete a separate questionnaire focused on the teacher application. The questionnaire mirrored the structure of the student version, with sections adapted to reflect the specific features and workflow of the teacher interface.

The application-specific section was adapted to reflect the teacher's interaction flow, with items on a 5-point scale addressing the clarity of the system tutorial provided at the start of the session, the ease of planning an activity, the comprehensibility of the overall workflow from activity planning through to the visualization of results, and the clarity of the Results panel. A multiple-choice item allowed participants to identify activities whose results were unclear, with an optional open-ended follow-up. As in the student questionnaire, dedicated fields were included for the reporting of technical issues, and optional open-ended fields collected general impressions and suggestions regarding the teacher application.

The full list of questionnaire items is reported in Table 2.

Table 2: Teacher questionnaire items

Section	Items
Demographic Information	1. Age range 2. Gender 3. Occupation 4. Background 5. Which of the three sessions did you attend as a teacher? (single choice)
System Usability Scale (SUS) [?]	1. I think that I would like to use this system frequently. 2. I found the system unnecessarily complex. 3. I thought the system was easy to use. 4. I think that I would need the support of a technical person to be able to use this system. 5. I found the various functions in this system were well integrated. 6. I thought there was too much inconsistency in this system. 7. I would imagine that most people would learn to use this system very quickly. 8. I found the system very cumbersome to use. 9. I felt very confident using the system. 10. I needed to learn a lot of things before I could get going with this system.
User Experience Questionnaire – Short Version (UEQ-S) [2]	1. Obstructive – Supportive 2. Complicated – Easy 3. Inefficient – Efficient 4. Confusing – Clear 5. Boring – Exciting 6. Not interesting – Interesting 7. Conventional – Original 8. Usual – Modern
App-Specific Usability Items	1. How clear was the tutorial provided by the system? 2. How easy was it to plan an activity? 3. How clear was the full flow between planning, managing planned activities, launching activities, and viewing results? 4. How clear did you find the Results panel, considering all activities? 5. Was there any activity whose results were particularly unclear? (multiple choice) 6. If so, what issues did you encounter? (open-ended, optional)

Section	Items
Technical Issues	1. Did you encounter any technical bugs during use? (open-ended, optional) 2. Did you experience any stability or performance issues? (open-ended, optional)
Open-Ended Feedback	1. What struck you most about this system? (optional) 2. What would you change in the teacher app? (optional) 3. Any other comments or general suggestions? (optional)

3 Participants

The evaluation was conducted across two independent sessions held in two different locations in Italy during Summer 2025: the first in **Cesena**, Emilia-Romagna, and the second in **Termini Imerese**, Sicily. Both sessions were facilitated by members of the project team, who attended in the role of observers and moderators. Participants were recruited through direct contact and snowball sampling. Overall, **13 participants** took part in the evaluation, 6 in the first session and 7 in the second. The group included 8 male participants, 4 female, and 1 who preferred not to disclose their gender. In terms of age, 9 participants fell within the 22–25 range and the remaining 4 within the 26–29 range. Participants represented a diverse range of educational backgrounds and occupational profiles. Regarding occupation, the group comprised 2 bachelor's students, 4 master's students, 4 PhD students, and 3 workers (2 self-employed and 1 employee). In terms of disciplinary background, participants came from a variety of fields, including informatics (5), medical studies such as pharmacy and nursing (3), education (1), design (1), economics (1), physiotherapy (1), and sport sciences (1). Among the 13 participants, **6 volunteered to take on the role of teacher** during the sessions, one per session as per the evaluation protocol. This subgroup consisted of 4 male and 2 female participants, with 3 aged 22–25 and 3 aged 26–29. In terms of occupation, it included 1 bachelor's student, 1 master's student, 3 PhD students, and 1 self-employed worker. Their disciplinary backgrounds were evenly split between informatics (3) and medical studies (3). A summary of participants' demographic and background characteristics is reported in Table 3. All participants completed the questionnaire concearning the student

Table 3: Demographic and background characteristics of participants

Characteristic	All participants (N=13)	Teachers (N=6)
Gender		
Male	8	4
Female	4	2
Prefer not to answer	1	0
Age range		
22–25	9	3
26–29	4	3
Occupation		

Characteristic	All participants (N=13)	Teachers (N=6)
Bachelor's student	2	1
Master's student	4	1
PhD student	4	3
Self-employed worker	2	1
Employee	1	0
Disciplinary background		
Informatics	5	3
Medical studies	3	3
Education	1	0
Design	1	0
Economics	1	0
Physiotherapy	1	0
Sport sciences	1	0

4 Results

This section presents the findings emerging from the first evaluation of the platform, conducted across two in-the-lab sessions with a total of 13 participants. Given the exploratory nature of this evaluation and the limited sample size, the analysis does not aim to draw statistically generalizable conclusions, but rather to provide an initial, evidence-based picture of the platform's strengths and areas for improvement from the perspective of both user roles. Results are presented separately for the **student application**, based on the responses of all 13 participants, and for the **teacher application**, based on the responses of the 6 participants who also took on the teacher role. For each application, quantitative data derived from the SUS and UEQ-S scales and the application-specific items are discussed alongside the qualitative feedback collected through open-ended responses, in order to provide a comprehensive and nuanced account of participants' experiences with the platform.

4.0.1 System Usability Scale (SUS)

The SUS scores obtained from the 13 participants who evaluated the student application indicate a generally positive perception of its usability. The mean SUS score was **81.92** (SD=12.10), which falls within the *Excellent* range according to the standard SUS grading scale [1], and comfortably exceeds the commonly adopted threshold of 68 points that distinguishes above-average from below-average usability. The distribution of individual scores, illustrated in Figure 1, further confirms the overall positive assessment, while also highlighting some degree of variability among participants, suggesting that certain aspects of the interface may have been perceived differently across users with different backgrounds and levels of familiarity with similar platforms.

The SUS scores obtained from the 6 participants who evaluated the teacher application similarly reflect a positive usability assessment. The mean SUS score was **80.00** (SD=16.39), also

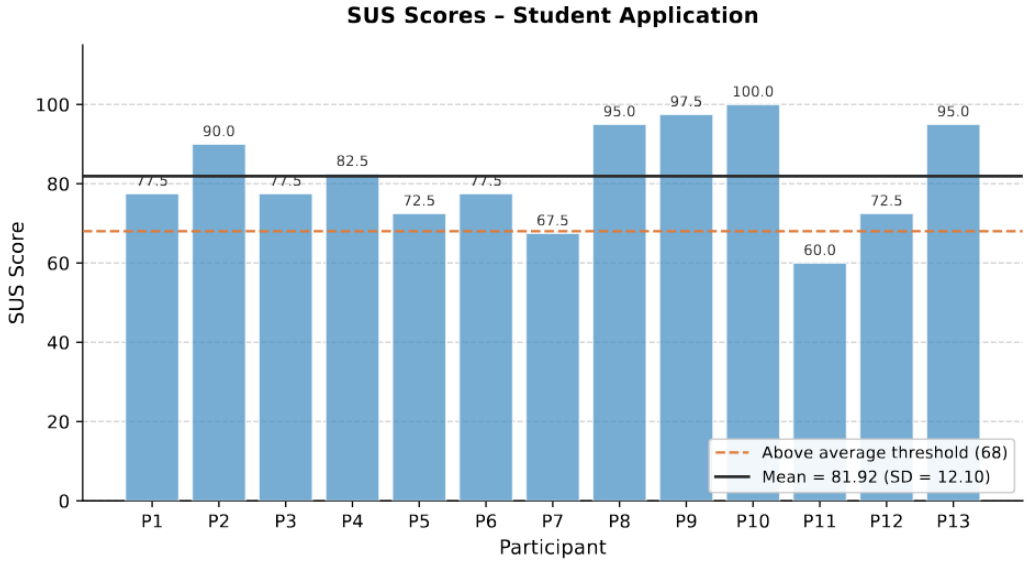

Figure 1: Distribution of individual SUS scores for the student application. The dashed line indicates the mean score (81.92), and the dotted line marks the standard above-average threshold (68).

falling within the *Excellent* range of the SUS grading scale [1] and well above the 68-point threshold. The higher standard deviation observed in this subgroup, compared to the student application, should be interpreted with caution given the small sample size; nonetheless, it suggests a somewhat wider spread in individual perceptions. The distribution of individual scores is shown in Figure 2.

4.0.2 User Experience Questionnaire – Short Version (UEQ-S)

The UEQ-S results provide an initial overview of participants' perceptions of the platform's user experience, considering both pragmatic and hedonic aspects. For the **student application**, the mean scores were **1.56** (SD = 0.82) for *Pragmatic Quality* and **1.10** (SD = 1.33) for *Hedonic Quality*. These values suggest that students perceived the application as generally usable and reasonably engaging, with slightly lower scores on hedonic dimensions, indicating that novelty and stimulation may be areas for potential enhancement.

For the **teacher application**, the mean UEQ-S scores were **1.50** (SD = 1.03) for *Pragmatic Quality* and **1.38** (SD = 1.07) for *Hedonic Quality*. These results indicate that teachers also experienced the platform positively, with both usability and hedonic appeal rated favorably. The slightly higher hedonic score compared to students suggests that teachers found the platform more stimulating or enjoyable, despite the small sample size (N = 6).

Figure 3 and Figure 4 illustrate the distribution of UEQ-S scores across participants for each application, providing a visual summary of both pragmatic and hedonic evaluations.

4.0.3 App-Specific Usability Items

Table 4 reports the mean scores and standard deviations for the app-specific usability items related to the student application (1 = not clear at all, 5 = very clear). Overall, the results

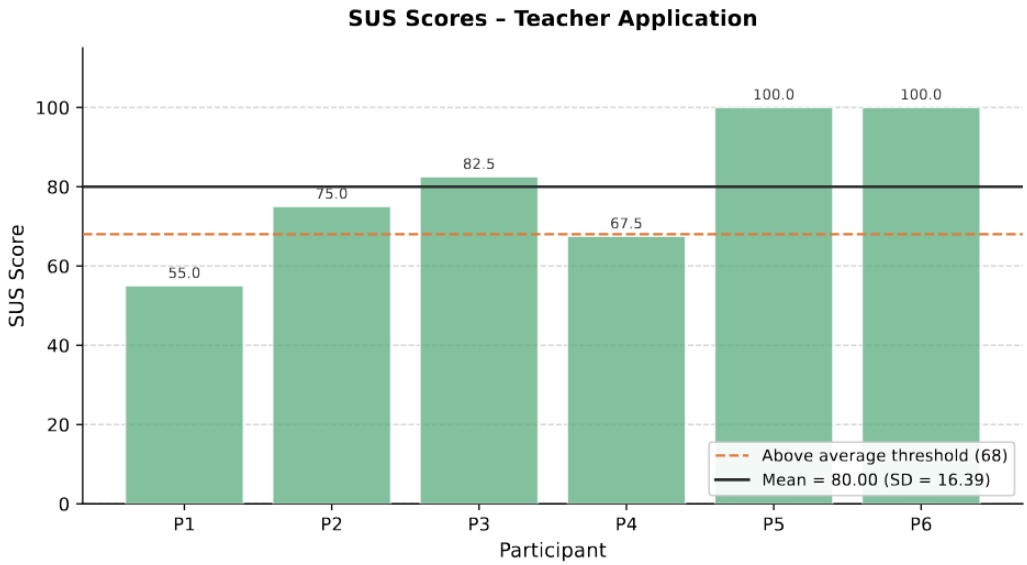

Figure 2: Distribution of individual SUS scores for the teacher application. The dashed line indicates the mean score (80.00), and the dotted line marks the standard above-average threshold (68).

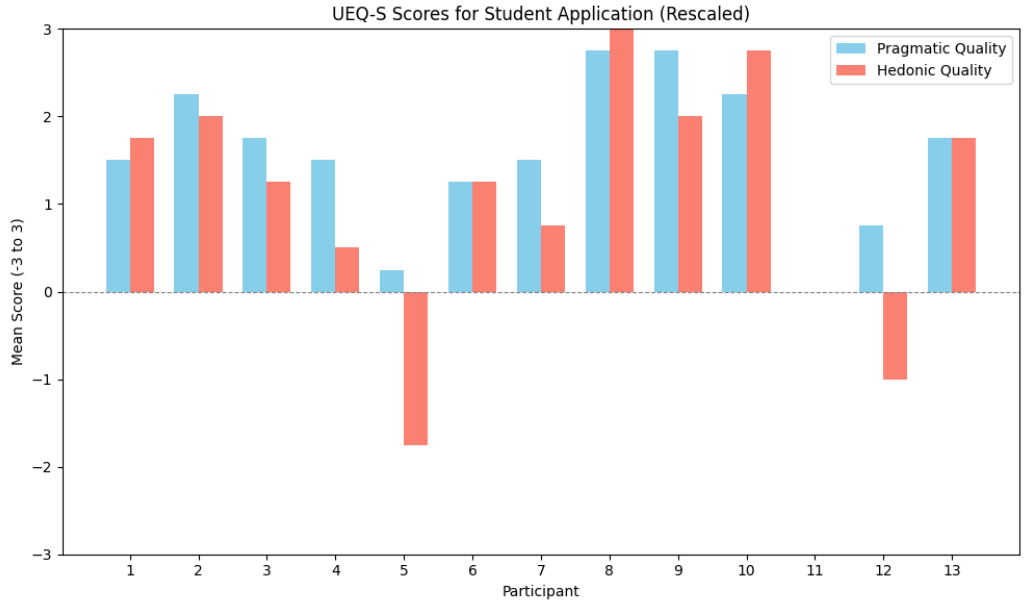

Figure 3: UEQ-S scores for the student application: mean scores for Pragmatic and Hedonic Quality (N = 13).

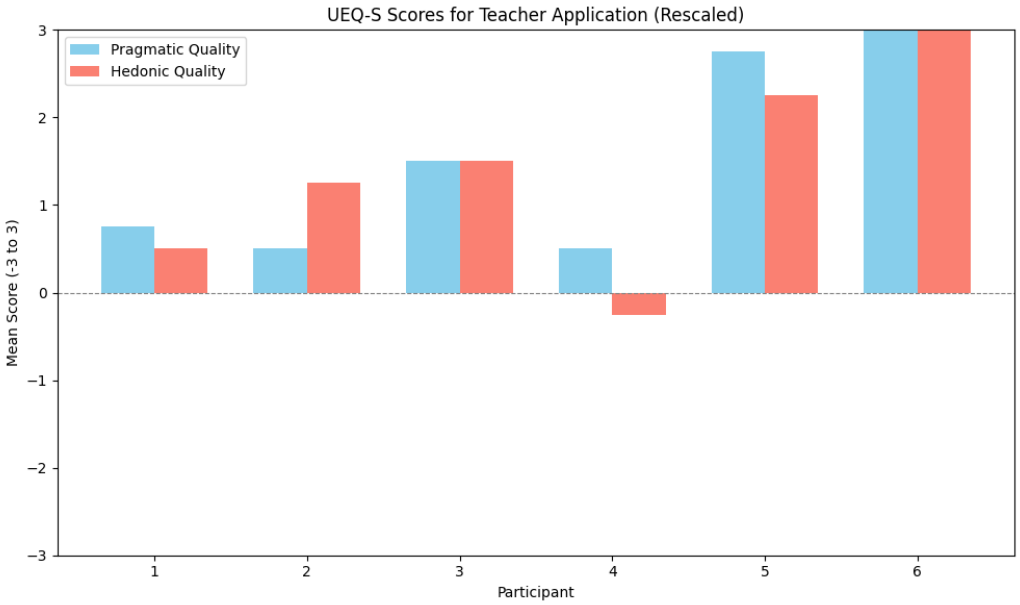

Figure 4: UEQ-S scores for the teacher application: mean scores for Pragmatic and Hedonic Quality (N = 6).

indicate that students perceived the application as generally clear and easy to use across the evaluated dimensions.

The highest score was observed for *Ease of completing activities* (M = 4.46, SD = 0.52), suggesting that most participants were able to navigate and complete the activities without major difficulties. Similarly, both the *Clarity of the activity flow* (M = 4.08, SD = 0.95) and the *Clarity of gamification elements* (M = 4.08, SD = 1.04) received positive evaluations, indicating that the structure of the application and the gamified features were generally understandable.

The comparatively lowest score was observed for the *Clarity of the Results panel* (M = 3.85, SD = 1.14). While still above the midpoint of the scale, the larger variability suggests that some students experienced difficulties interpreting the feedback and results provided by the system. This aspect therefore represents a potential area for improvement in future iterations of the platform.

Item	Mean	SD
Ease of completing activities	4.46	0.50
Clarity of activity flow	4.08	0.94
Clarity of gamification elements	4.08	0.95
Clarity of 'Results' panel	3.85	0.99

Table 4: Mean scores and standard deviations for student app-specific usability items (N=13).

Table 5 reports the results for the teacher-facing application. Overall, teachers also reported positive perceptions of clarity and usability across the evaluated items.

The highest score was observed for *Ease of planning activities* (M = 4.33, SD = 0.82), indicating that the process of creating and scheduling activities was perceived as straightforward. The

Clarity of the tutorial (M = 4.00, SD = 0.63) and the *Clarity of the activity flow* (M = 4.00, SD = 0.63) also received positive evaluations, suggesting that teachers were generally able to understand how to interact with the system after the initial guidance.

The *Clarity of the Results panel* received a slightly higher mean score (M = 4.17, SD = 0.75), indicating that teachers generally found the results view understandable. However, qualitative feedback (reported in the following subsection) revealed that some participants would prefer more detailed information about individual student responses and improved visualization of results.

Overall, these findings suggest that both the student and teacher applications provide a generally clear and usable experience, while highlighting specific components—particularly the presentation of results—that could benefit from further refinement.

Item	Mean	SD
Clarity of tutorial	4.00	0.63
Ease of planning activities	4.33	0.82
Clarity of activity flow	4.00	0.63
Clarity of 'Results' panel	4.17	0.75

Table 5: Mean scores and standard deviations for teacher app-specific usability items (N=6).

4.0.4 Qualitative Feedback

In addition to the quantitative measures, participants were invited to provide open-ended feedback regarding their experience with the platform. The responses were analyzed through a lightweight thematic grouping in order to identify recurring observations, usability issues, and suggestions for improvement. Although the sample size was limited, several themes emerged consistently across participants.

Many students highlighted the overall *simplicity of use* and the *originality of the platform*. In particular, the gamification elements (e.g., badges, coins, profile customization) were frequently mentioned as engaging features that made the experience more motivating. Several participants also appreciated the *visual design and responsiveness* of the application, reporting that activities were generally easy to access and complete.

Teachers similarly emphasized the *speed and ease of planning activities* as a positive aspect of the system. The ability to quickly schedule activities and manage sessions was perceived as particularly valuable for classroom use.

While most students reported that the activities were generally understandable, some participants noted difficulties in interpreting the *results or feedback provided after completing certain activities*. In particular, a few students reported that it was sometimes unclear when an activity was fully completed or how the results should be interpreted.

Teachers also expressed interest in having more detailed information in the results views, such as clearer visualizations or the possibility to inspect individual student responses more easily. These comments suggest that improvements in the presentation and explanation of results could enhance the overall usability of the platform.

Several participants proposed improvements to the interface and interaction design. For example, students suggested introducing *additional animations and feedback* for gamified elements (e.g., animated growth in the "forest" activity or clearer notifications when unlocking trophies). Other suggestions included improving the *layout spacing between interface elements*, enhancing the *dark mode color contrast*, and adding *notifications* to signal important events such as the start of an activity or the unlocking of rewards.

Teachers proposed improvements to the tutorial and navigation, including clearer guidance when performing actions and more contextual explanations during the onboarding tutorial.

Several participants reported technical issues encountered during the testing sessions, including occasional application crashes, rendering errors, interface elements not displaying correctly on some devices, and temporary difficulties loading activities or results. Some students also reported issues related to dark mode visualization, font scaling on larger displays, and unexpected logouts when switching between applications.

In addition to the feedback collected through the questionnaires, the session facilitator recorded observations and notes during the testing sessions. These notes included additional instances of bugs, unexpected behaviors, and moments in which participants appeared uncertain about how to proceed within the application. These observations were used to complement the self-reported feedback and helped identify specific interface elements and workflows that require further refinement.

Overall, the qualitative feedback provided actionable insights that informed future iterations of both the student and teacher applications, particularly regarding the clarity of results visualization, interface feedback mechanisms, and the resolution of technical issues identified during testing.

5 Conclusion

The results of this evaluation provide valuable insights into the usability, functionality, and overall user experience of the proposed digital wellbeing platform, considering both the student and teacher applications. By combining quantitative usability measures, qualitative feedback from participants, and direct observations recorded during the testing sessions, the study offers a comprehensive understanding of how users interact with the system and which aspects require further refinement.

Overall, the evaluation highlighted several positive aspects of the platform. Both students and teachers generally perceived the system as usable and understandable, and many participants appreciated its originality, the integration of reflective activities on digital wellbeing, and the presence of gamification elements aimed at increasing engagement. Students particularly valued the visual design, responsiveness, and gamified features such as avatars, badges, and the virtual shop, while teachers emphasized the usefulness of the activity planning functionality and the variety of educational activities available in the system.

At the same time, the testing sessions revealed a number of technical issues and usability limitations that guided the development of future iterations of the platform.

Regarding the teacher application, the usability questionnaire indicated that some participants perceived the system as somewhat confusing and not always easy to use. Qualitative feedback and facilitator observations suggested that this perception may be related to the complexity of certain workflows and the amount of textual information present in the interface. Several improvements therefore emerged as priorities for future versions of the application. These include simplifying the interface and reducing excessive textual content, introducing clearer navigation mechanisms and improving the onboarding tutorial by highlighting the interface elements being explained. From a functional perspective, teachers expressed interest in additional features for monitoring classroom activities, such as clearer visualizations of student responses, the possibility to inspect contributions in word clouds, and indicators showing whether all students have completed an activity. Furthermore, several technical issues were identified, including graphical bugs in the bar chart visualizations, inconsistencies in the word cloud results, errors in the brainstorming activity, issues related to the forest activity planning interface, and compatibility problems with certain browsers such as Firefox. Addressing these issues will be essential to ensure reliability and consistency of the teacher interface.

Similarly, the student application revealed several areas for improvement, primarily related to technical stability and clarity of interaction flows. During the testing sessions, multiple bugs were identified across different devices and operating systems. These included crashes in specific activities (such as the bingo and brainstorming tasks), inconsistencies in the management of virtual currency and trophies, occasional failures in loading images in storytelling activities, and login issues on certain iOS devices. Additional technical problems included unexpected logout events, interface rendering issues in dark mode, and difficulties related to automatic screen refresh mechanisms, which sometimes caused disruptive visual flashes or forced the user back to the top of the page.

Beyond the technical aspects, the feedback collected from students highlighted several opportunities to improve the user experience and the gamification design. Participants suggested making certain activities clearer, particularly the bingo activity and the feedback shown at the end of tasks. Several students also expressed interest in stronger feedback mechanisms, such as clearer notifications when new activities start or when trophies are unlocked, visual indicators showing unseen achievements, and small rewards associated with trophies to reinforce the gamification dynamics. Other suggestions included improving the contrast between text and background images in storytelling sections, introducing animations for elements such as the forest activity, and making it more explicit when an activity has been completed.

The facilitator observations recorded during the testing sessions provided an additional source of information that complemented the questionnaire responses. In several cases, users showed hesitation or uncertainty when navigating specific interface elements or interpreting activity results, confirming some of the issues reported in the qualitative feedback. These observations were particularly useful for identifying moments of confusion in the interaction flow and for detecting bugs that were not always explicitly reported by participants.

Taken together, the findings of this evaluation provide a clear roadmap for the subsequent development iteration of the platform. To improve the application we focused on improving the technical stability of both applications, refining the clarity of interaction flows, enhancing the feedback mechanisms associated with gamification elements, and simplifying certain as-

pects of the user interface to reduce cognitive load. Additional improvements also targeted the visualization of activity results and the monitoring tools available to teachers.

Despite the identified limitations, the evaluation confirms the feasibility and potential value of the platform as a tool for supporting educational activities related to digital wellbeing. With the implementation of the improvements identified during this study and further rounds of testing, the system had the potentiality to evolve into a more robust and engaging solution capable of effectively supporting both students and teachers in promoting more conscious and balanced technology use.

References

[1] Aaron Bangor, Philip Kortum, and James Miller. Determining what individual sus scores mean: Adding an adjective rating scale. *Journal of usability studies*, 4(3):114–123, 2009.

[2] Andreas Hinderks. Design and evaluation of a short version of the user experience questionnaire (ueq-s). *International Journal of Interactive Multimedia and Artificial Intelligence*, 2017.

[3] James R Lewis. The system usability scale: past, present, and future. *International Journal of Human–Computer Interaction*, 34(7):577–590, 2018.