User Evaluation Report

Cohort 1, Group 6 - M6

Members:
Mir Baydemir
Adam Fraulo
Azib Hj Abu Akmar
Agata Pittarello
Esther Scanlon
Jazz Stubbins
Sonia Szetela

User Evaluation Method:

Recruitment[1,2]:
Three participants were recruited from other teams within cohort 1.I made sure that the participants were over 18 years of age and had experienced university exam challenges. All participants had relevant gaming and technical experience as they had developed similar systems themselves. This approach aligns with the task based user evaluation principle of participants in the user evaluation should be similar to the final users of the system.
Each participant was provided with an information sheet before the session began. Participants completed their online consent forms before their sessions began, confirming that their participation is voluntary and understanding that the evaluation assesses the prototype system rather than their abilities.

Data Collection Tools and Data:
I have designed 9 tasks to test game functionality:[3] Launch the game and navigating menus(Task1), reading instructions and understanding objectives of the game(Task 2),starting the game and initial impressions(Task 3), testing game controls simplicity and testing the maze environment(Task 4), pause functionality(Task 5), interacting with visible negative events(Task 6), interacting with visible positive events(Task 7), finish the game in 5 minutes(Task 8), view leaderboard with top 5 scores(Task 9). This progressed logically from simple to complex tasks to build participant confidence before testing critical functionality. Interview response sheets were created using Google Docs. These sheets captured data through observation checklists(yes/no responses) documenting completion success of tasks, feature visibility and whether the help is requested or not. Qualitative data was collected by asking participants to speak their thoughts aloud as they worked through tasks. This technique is known as the think aloud protocol. I recorded usability issues with severity ratings on a 1-4 scale.After the test, I asked participants some post-test questions to gather overall experience ratings on a scale of 1-5. In addition, some of the post-test questions were open-ended asking about likes , dislikes and improvement suggestions.

Procedures:
I conducted every evaluation session face to face with participants. After completing the consent forms, I asked participants warm-up questions to prepare them for the main tasks.They then performed nine tasks sequentially while encouraged to verbalize their thought processes which enabled me to understand their mental models of the system. Two team members worked together during each session: one as interviewer and one as observer. I as an interviewer asked questions and guided participants, while the observer recorded their responses and behaviours on the interview sheets. When participants encountered difficulties, I encouraged them to explain their problem, being careful not to ask them leading questions that might influence their responses. I made clear to participants that any problems they encountered were due to issues with system design not their abilities. Each session lasted approximately 30 minutes. To protect participants anonymity, we did not audio or video record the sessions and participant names were not included in the interview sheets.

Usability Problem Table:

Key of severity ranking:
1-Cosmetic problem
2-Minor problem
3- Major problem
4- Catastrophic problem

Usability Problem	**Participants Reporting**	**Severity Ranking (1-5)**	**Description**
Instructions not visible on initial menu launch	P3	4 (catastrophic)	Menu did not display instructions upon launch, which prevented the user from accessing the game instructions.(bug)
Leaderboard missing	P1 P2 P3	4 (catastrophic)	The leaderboard feature was absent from the game, which made it impossible to view top scores as required by game objectives.
No replay button after completing the game	P2	3 (major)	After the game ended, there was no option to replay without relaunching the application.
Insufficient instructions for understanding controls	P2	2 (minor)	Instructions did not clearly explain what the ‘E’ key does, leading to confusion during the game.
Unclear map boundaries	P2	3 (major)	Difficult to distinguish between explorable areas and boundaries, especially corridor areas.
Unclear exit location	P2	3 (major)	The player couldn’t find the exit which makes the primary objective unclear.
Positive events difficult to identify	P2 P3	2 (minor)	Players were unsure which items were positive events therefore guidance was required.
Negative events difficult to identify	P2	2 (minor)	The player only recognized the dean as a negative event, and couldn't identify others.

References:
[1] Vasiliou, C. (2025) User Evaluation , Engineering 1: Software & Systems Engineering 1, Week 11 [Lecture]. University of York Virtual Learning Environment (VLE). (Accessed: December 2025)

[2] Sharp, H. , Preece, J. and Rogers, Y. (2023). ‘Evaluation studies: from controlled to natural settings’ , in Interaction Design : Beyond Human Computer Interaction. John Wiley & Sons, Chapter 15

[3] User Evaluation Sheet: https://keybordkat.github.io/websiteGroup6/User%20Evaluation%20Interview%20Sheet.pdf