Jungle Riddles: Traversing the Temple's Trials

An Interactive Story

Supervisor: Charlotte Stinkeste (csti@kth.se)

Team:
Shashank Shirol (shirol@kth.se)
Veerle Uhl (veerleu@kth.se)
Elisabetta De Luca (elidl@kth.se)
María Osorio Vergara(mcov@kth.se)

Video Demo:
https://vimeo.com/902270834

Abstract

This report describes the development and evaluation of a game that uses the phone as a multimodal controller for an interactive storytelling application. It uses a socket.io server-client setup to allow the phone to interact with a PlayCanvas game instance built to include different modalities like blowing, ambient light detection, and tilting. User testing concluded that the multimodal setup resulted in high levels of immersion and enjoyment for the users. The phone as a multimodal controller can thus be seen as a cheap alternative for games to include interaction as a way to increase user immersion in the story.

Introduction & previous work

Ever since we existed, humans have loved to share stories. Even before we could write, we would pass on stories through word-of-mouth, transforming them into songs even to make them more easily remembered. Nowadays, traditional storytelling has made its way from books to movies, series, and even video games. Some video games have become so story driven that the Game Awards even have a category of awards they hand out each year for best narrative in a video game (The Game Awards, n.d.). But what makes a narrative the “best”? Which storytelling can be seen as good, and which as bad? Immersion of games and stories is often a category looked at for this. When a user can be immersed in a story, they will feel like they are really part of the story, or feel like their actions shape the outcome of the story. Roth, P. C. H. (2016) differentiates between two different forms of immersion in games. First, immersion can take place on a local level of specific scenes. This means that when a user interacts with objects in a scene, the system recognizes this and shows cause and effect of a user's actions on the environment. The second form of immersion takes place on the global level of narrative. Here, users can interact with elements of the story, which in turn decide, and perhaps change, the outcome of the story. Unfortunately, Roth did not find a significant improvement in enjoyment for users that were able to change the narrative through their actions, and users who weren’t.

Cavazza et al. (2007) also did a study where users could interact with characters through a digital platform. In their study, Cavazza transformed an old classic story, *Madame Bovary,* into a VR environment, where users could interact with Madame Bovary through speech. In their research, they use the CAVE system as a way to get a user virtually immersed into the story. The researcher discovered that having a user be able to constantly interact with other characters in the story caused much strain on the user for having to constantly focus on what to say. Moreover, a VR implementation such as this is not likely to be feasible for many average consumers to use, since it is a costly and big system to put in place.

In this project, we want to test what the different options are to use as a low-cost alternative for fully immersive VR settings, where a user would still be able to interact with the story following the first level of immersion as defined by Roth. In this study the goal is to focus on overall user experience, and to compare the different modalities as forms of immersive interaction with the storytelling. In order to add the interactive element to the final project, we will allow users to open up a website on their phone, and respectively use the sensor from the phone as inputs for the storytelling application, which will run on a users' laptop.

There have been many projects on the market in the last few years all exploring different kinds of digital interfaces as inputs for games. The rise of smartphone games and game controllers has opened up many new modalities for gaming, such as haptic feedback, tilt controls, and even touchscreen controls. Valente (2008) started exploring different phone sensors in games more than 15 years ago. They explored the use of sound and tilt controls as a way to make a "blind" game on phones, where no visual feedback was given at all. Vajk et al. (2008) integrated the different phone sensors for games differently. They explored the use of the phone and its tilt sensors as a way to control movement on a second device (a big screen or a monitor), connected through Bluetooth. They used the phone in a Wii-like way, where the phone's tilt sensors acted similar to a Wii controller. Huang, Lu, and Fang (2011) took this a step further and used a phone as a way to control a computer game. They mapped the controls that were normally evoked through mouse and keyboard, to a phone that a player could control. The phone and the computer were connected through WiFi. Players could control the game through a touch screen, microphone, and G-sensor. The phone then gave feedback back to them through its speaker and vibrations.

The project

Figure 1: *A screenshot of the starting scene of the game.*

The final product implementation

After getting the initial idea of making a storytelling game that combined both phone sensor input and visual output on a computer screen, the first step was exploring the

various ways in which a connection can be made between two different interfaces. After some exploration, it was found that socket.io(n.d.) was the best fit for this design, allowing bidirectional and low-latency communication for every platform. Using socket.io, we set up a server-client paradigm where the application running on the computer (server) polls the sensor data from the phone (client) and sends to a PlayCanvas instance (client). There are triggers in place that transmit data (sensor and other data) once a condition is satisfied.
The socket.io server is run on Glitch, which is an online environment to build and host full-stack web apps. We set up the server on Glitch that serves various pages that correspond to different puzzles in the game. We use PlayCanvas as a way to animate the storytelling part on the web platform. In addition to the glitch server, we also implement a Glitch client, which runs as a Glitch website on a user's phone. This way, both the phone, using Glitch, and the computer storytelling, using PlayCanvas, would be able to communicate with each other as two clients, sending data to and receiving data from the Glitch server.

In the PlayCanvas setup, assets from various online free sources were used to decorate the 3D environment in which the story takes place. In order to convey the story elements, text is shown on screen in which users can read about the background story, as well as read the hints to solve the puzzles. The PlayCanvas application has different scenes for the different puzzles. Once a scene is rendered and displayed on the PlayCanvas instance, it will send a signal to the Glitch website which will then push the right html file to the phone. Thus, for each puzzle in the game, there is an accompanying PlayCanvas scene and Glitch page. Once a puzzle scene is rendered and displayed with PlayCanvas, the PlayCanvas instance will listen for input received from the Glitch server, which in turn gets its multimodal input from the different interactions a user will perform with their phone. Once the correct input is registered, the puzzle scene plays out accordingly; after a puzzle is solved, Glitch will trigger rendering of the next puzzle which is shown in the same PlayCanvas instance.

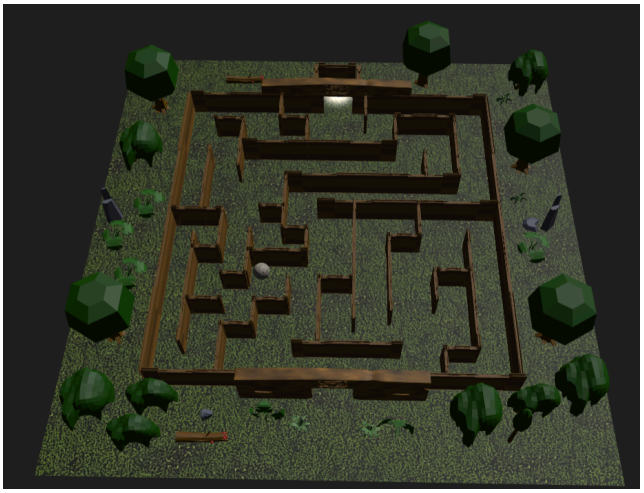
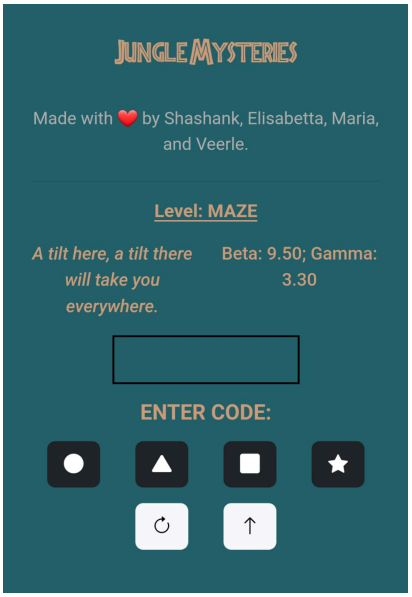
1)  2) 

Figure 2: *Snapshots of the game during the second level. 1) A screenshot of the desktop game in the maze level. 2) A screenshot of the phone game during the maze level.*

Multimodal and HCI implementation

The final application makes use of different modalities in order to create immersion for the user. As inputs, users can tilt their phone, blow into the microphone of their phone, cover the ambient light sensor, or press on the screen. All these different input modalities are mapped to different actions which translate the movements they are doing with their phone to movements in the game. In order for immersion to appear, the actions are closely mapped to their outcomes. For example, blowing into the microphone makes bushes blow away in the story, thus an action of the user directly translates into an action in the game. In this way, it is intuitive for a user to understand the consequences of their actions. As output modalities, both sound and visuals are used to strengthen this feeling of immersion and assist in showing the outcomes of a user's actions on their phone. We also implement haptic feedback to alert the user about wrong attempts when they try to solve a puzzle.

One limitation of using a laptop or computer screen as output modality is that it will result in less immersion, compared to an AR or even VR system. However, the computer method does allow for a more accessible system that can be run by almost everyone, compared to a pricey VR system which not many people have access to.

User testing

As said in the introduction, the aim of this research is to discover whether the different modalities used by the interaction of a phone and computer application add different levels of immersion and user experience to a storytelling application. In

order to test how different modalities influence user experience and immersion, the final application has multiple modalities implemented into the game. During the user testing, the goal is to discover if a user has different reactions to the different “minigames”.

User test setup

During the user test, a user is given a phone with the Glitch website opened, and sat in front of a laptop with the PlayCanvas open. Before starting the application, the user is informed of the risks of the user testing (which are none), and informed of the general outline of the test. Users are informed that different modalities will be tested as a way of doing minigames, but are not told what factors they will be evaluated on. Next, the storytelling application will start, and users will play the game. Experimenters will not interfere during this process, but will observe and write down their observations on actions that the user performs. It was opted for silent observation instead of a thinking aloud procedure by the user, since a thinking aloud from the user could break the immersion that would be measured.

After the user finishes playing the game, they will get a final questionnaire. In this questionnaire, questions are asked from the ARI immersion scale (Georgiou, Y., & Kyza, E. A., 2017) to measure engagement, engrossment, and total immersion in the application. These scales have Cronbachs alpha’s of 0.77, 0.88, and 0.82 respectively. For questions where an AR system is described, this was replaced with “the application”. The questions asked in this questionnaire are presented in Appendix A.

User test results

Since there was only a limited timeframe available for this project, a limited user study was performed. A total of 3 participants were used for the user test, ranging from 20 to 50. This is not the ideal amount of participants needed for a reliable and representative user test, but due to time constraints there was no other option. The decision was made against testing with the target group (small children) since access to that target group was sparse for the researchers and more guidelines would have to be passed for that to be possible.

Let us take a look at the results from the survey. In the survey, questions about engagement, engrossment, and total immersion were asked. In figure 3 the results of the survey are displayed. In general, levels in all categories are above average, indicating immersion to some degree.

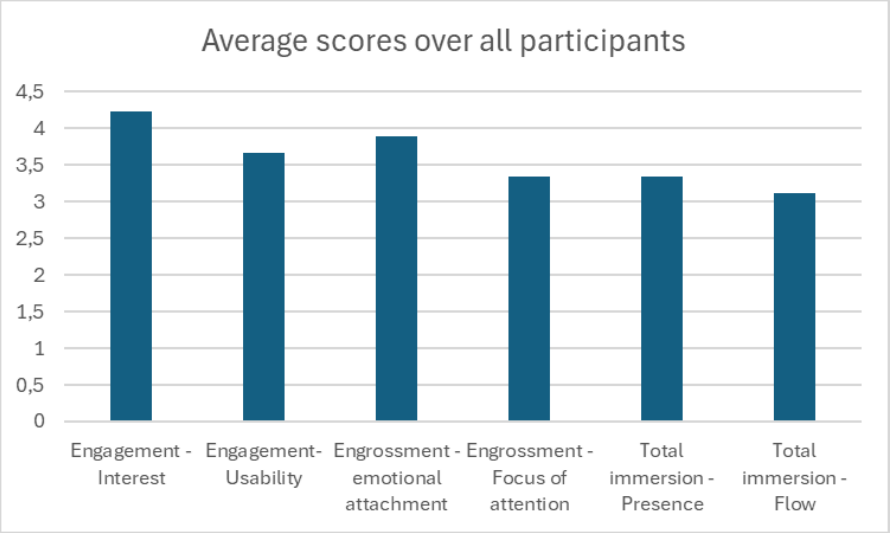


Figure 3: *Average scores from the user test survey*

Engagement

Figure 4 below shows the results from the survey regarding the questions about interest. In this figure, the first three questions relate to interest while the last four questions relate to usability. Note that question 6 and 7 are inversely coded. From this graph it is clear that people rated the game quite highly on interest, and thus were quite interested in playing the game. The usability questions show that the ease of playing was moderately high. Only one participant found the game confusing and none of the participants found the game complex.

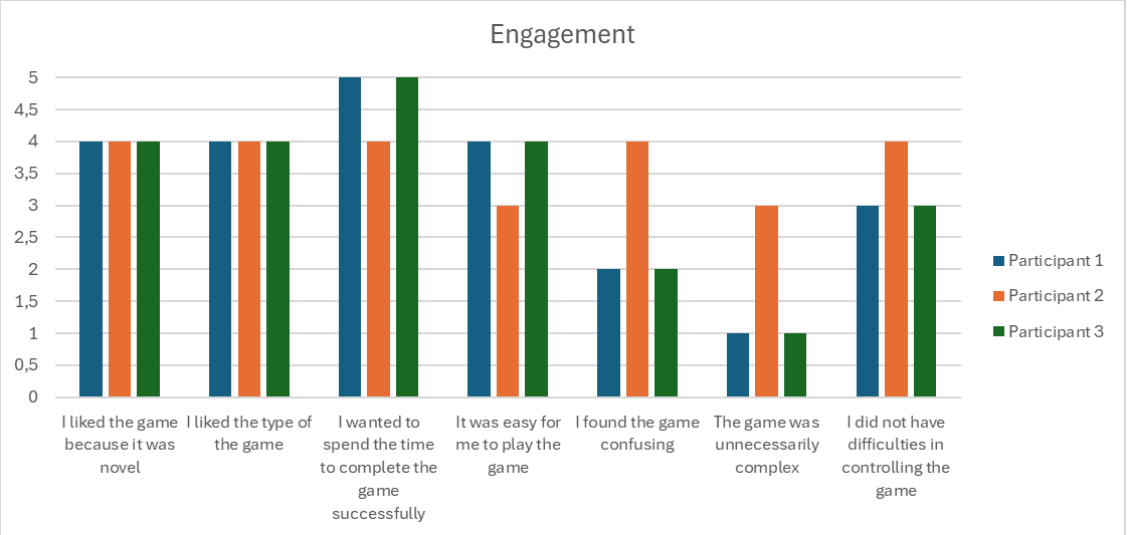


Figure 4: *Results from the engagement questions in the user survey*

Engrossment

In figure 5 the survey results for the engrossment category are displayed. The first three questions here relate to the emotional attachment of users, while the last two questions relate to the focus of attention. It is clear that emotional attachment is really quite high for all participants. Users were curious about how the game would progress and were excited to play it. Focus is moderately high for all participants.

Users rated an average of 3.33 on focus of attention during the game. This lack of focus could be explained by the fact that a researcher present during the test can be distracting to the users.

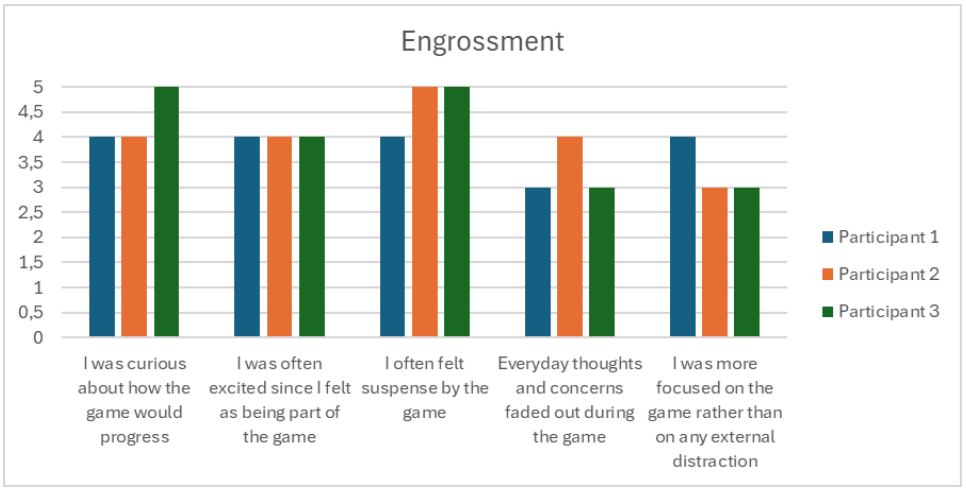


Figure 5: *Results from the engrossment questions in the user survey*

Total immersion

Figure 6 shows the results of the questions related to total immersion. Here, the first four questions relate to presence, while the last three relate to flow. Presence scores for all participants were close to neutral, but an interesting deviation from this is the question “I was so involved in the game, that in some cases I wanted to interact with the virtual objects directly” where users answered with an average 4.3, indicating that local immersion was high for the participants. The flow score is relatively neutral for all participants. Participants indicate that their minds did wander and that they were able to think about other things while playing the game.

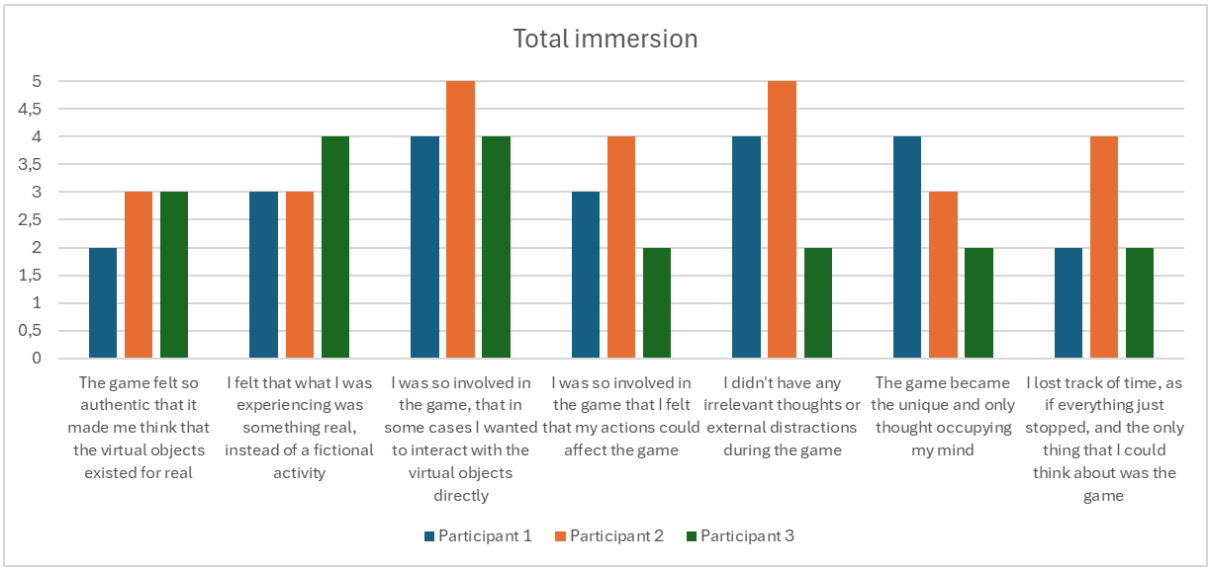


Figure 6: *Results from the immersion questions in the user survey*

Finally, experimenters also observed participants while they were performing the tasks and completing the game, and took notes of points that participants talked about. All participants indicated that they enjoyed the novel nature of the game, especially the use of different sensors on the phone. In congruence with the survey, participants also noted that they were curious about how the game progressed. Finally, the participants were asked to give some constructive feedback if they had any. Here, they noted that sometimes the symbols were a bit unclear, and that additional hints (especially at the start of the game) could be a nice addition. One participant commented on the large amounts of texts that they had to read through in order to understand the story, which the experimenters later improved by adding a text-to-speech module to read the story out loud.

Discussion and future work

At the start of this report, a research aim was set up. The goal of this research was to discover if a multimodal interactive storytelling application could result in feelings of immersion for the users. In order to reach this goal, an application was built which uses both a desktop and a mobile phone application to tell the story of Max, the jungle explorer, through interactive puzzles. In the introduction, different products were mentioned which, in some way, used controllers to interact with an application. This study took that notion of controllers one step further, and designed a controller which allowed for multimodal interaction with the one controller. For example, where the Wii controller only allows for button and tilt input, the phone controller also allows for light and microphone input.

After user testing, it was discovered that the application that was built did in fact cause the user to feel immersed in the story presented. Through the use of different modalities to solve the puzzles in the story, users felt that they were engaged in the story, sometimes even feeling as if they were really interacting with the elements displayed on screen. Users also commented on how the app was captivating to play, and overall a fun experience. The different modalities were seen as a novel approach to “standard” games and surprised the users, since they allowed for a new and interactive way of playing the game.

However, due to the limited time available for this project, there are areas on which the product could be improved upon. Firstly, due to the coding and setup of the sensor readings, only android phones can currently open and use the application. Moreover, before using the application, users will need to go to the Google Chrome flags to enable some flags that allow the application to read the sensor data properly. This makes it so currently the application cannot instantaneously be used by everyone, which is something that could be looked into for future work. Another point of improvement is related to the server-client setup. The way the application is set up now, there are two clients interacting with each other, and communicating through

the server. The server puts any information through from one client to the other, but does not check whether there is more than 1 client. As a result, it is not possible for 2 or more people to play the game at the same time. For the purposes of this study this was enough, but if the game were to be developed into a product to go on the market, then this problem would need to be solved, likely by adding some form of check to see which phone client corresponds to which desktop client. Finally, due to the nature of the sensors we are using as inputs, the environments in which the game can be played are somewhat restricted. For example, it would not be possible to play in a loud environment, since the blowing sensor would register the sounds as input. Additionally, a dark environment would immediately solve the first puzzle where a low light level needs to be detected.

In conclusion, the product designed in this study combined different previous research to build an interactive storytelling application which allows a user to interact with a story through different modalities, increasing the user's immersion in the story. It introduces the phone as a low-cost multimodal controller that has possibilities of expansion towards different modalities than researched in this report. The phone has many different sensors built in, and all of them can be used to allow users to interact with a game in more novel ways, increasing enjoyment and immersion.

References

Cavazza, M., Lugrin, J. L., Pizzi, D., & Charles, F. (2007, September). Madame bovary on the holodeck: immersive interactive storytelling. In *Proceedings of the 15th ACM international conference on Multimedia* (pp. 651-660).

Georgiou, Y., & Kyza, E. A. (2017). The development and validation of the ARI questionnaire: An instrument for measuring immersion in location-based augmented reality settings. *International Journal of Human-Computer Studies*, 98, 24-37.

Huang, W. C., Lu, T., & Fang, W. C. (2011, November). An innovative interface design with smart phone for interactive computer game applications. In *2011 IEEE International Games Innovation Conference (IGIC)* (pp. 62-65). IEEE.

Roth, P. C. H. (2016). *Experiencing interactive storytelling.*

Socket.IO. (n.d.). *Home*. Retrieved December 25, 2023, from https://socket.io/

The Game Awards. (n.d.). *The Game Awards Best Narrative Nominees 2023* https://thegameawards.com/nominees/best-narrative

Vajk, T., Coulton, P., Bamford, W., & Edwards, R. (2008). Using a mobile phone as a “wii-like” controller for playing games on a large public display. *International Journal of Computer Games Technology*, *2008*.

Valente, L., de Souza, C. S., & Feijó, B. (2008, October). An exploratory study on non-visual mobile phone interfaces for games. In *Proceedings of the VIII Brazilian Symposium on Human Factors in Computing Systems* (pp. 31-39).

Appendix A

User questionnaire on immersion

Engagement - Interest
A1: I liked the activity because it was novel
A2: I liked the type of the activity
A3: I wanted to spend the time to complete the activity successfully
A4: I wanted to spend time to participate in the activity
Engagement - Usability
A5: It was easy for me to use the application
A6: I found the application confusing *
A7: The application was unnecessarily complex*
A8: I did not have difficulties in controlling the application

Engrossment - Emotional attachment
B1: I was curious about how the activity would progress
B2: I was often excited since I felt as being part of the activity
B3: I often felt suspense by the activity
Engrossment - Focus of attention
B4: Everyday thoughts and concerns faded out during the activity
B5: I was more focused on the activity rather than on any external distraction

Total immersion - Presence
C1: The activity felt so authentic that it made me think that the virtual objects existed for real
C2: I felt that what I was experiencing was something real, instead of a fictional activity
C3: I was so involved in the activity, that in some cases I wanted to interact with the virtual objects directly
C4: I was so involved, that I felt that my actions could affect the activity
Total immersion - Flow
C5: I didn't have any irrelevant thoughts or external distractions during the activity
C6: The activity became the unique and only thought occupying my mind
C7: I lost track of time, as if everything just stopped, and the only thing that I could think about was the activity.

Feedback from the silent observation of the user playthroughs

User: Pasquale (64)
Observer: Betta
Date: 08/01/2024

- In the first puzzle hint, “show you the path” → suggest to use “way” instead of path
- The user was very confused with the X in the first code, he needed a bit of guiding to understand it was a multiplication
- In the second puzzle, the user expected to write the numbers directly, when it was clear he couldn’t write, he quickly understood the code
- The blowing puzzle works also with talking, the room was too loud and the bushes flew away without any action. Also the triangle was not immediate to grasp (the user thought it was a square at the beginning), but he got the code right the second try
- In general, it is not clear if the code is not working when the code is wrong as there is no user feedback

User: Angela (66)
Observer: Betta
Date: 08/01/2024

- The user needed plenty of time to understand the mechanism behind the first puzzle
- The user was very confused with the X in the first code, he needed a bit of guiding to understand it was a multiplication
- The confirm button was not very intuitive to find
- Maze and blow puzzle quite intuitive to grasp for the user and also easy to map the code in the first case and understand the symbol in the second

User: Aranza (21)
Observer: María
Date: 08/01/2024

- The user felt a bit overwhelmed by the amount of text on the screen, she felt that in the end it made the game lose her attention a bit.
- The user would prefer the voice telling the story to have pauses.
- The user did not immediately understand the first level code.
- The user asked if there was some kind of help, like a question mark on the screen to get more information on the level.
- The user wondered why the direction of the camera changes at each level.