UNIVERSITÉ DE FRIBOURG
UNIVERSITÄT FREIBURG

MASTER IN COMPUTER SCIENCE

TAP!TAP!

Multimodel User Interface

Supervisor: Prof.Denis Lalanne

Students: COSTA Kimmy, PATASHKEVICH Natallia, SWAMINATHAN Jennifer

May 26th, 2023

TABLE OF CONTENTS

1. Introduction

The development of innovative technologies has opened up new possibilities for creating inclusive solutions that assist individuals with visual impairments in their daily lives. In this project, our aim was to design and implement a multimodal system for color recognition targeting individuals with visual impairments and color blindness.

The concept of multimodality refers to the utilization of multiple sensory modalities to enhance the overall user experience and effectiveness of a system. In our case, we focused on incorporating speech and tangible modalities to create an interactive and engaging learning environment. By combining auditory and tactile cues, we aimed to provide individuals with visual impairments a more comprehensive and accessible means of learning color recognition.

The project involved the development of a gamified learning experience, where users could actively participate in color recognition exercises. By incorporating elements of gamification, such as interactive gameplay, we aimed to make the learning process enjoyable and motivating for the users.

The system utilizes an interactive table, two pucks, and a microphone for user interaction. To facilitate the multimodal learning experience, we utilized an interactive table as the central component of our system. Additionally, the system integrated audio capabilities for speech prompts and a visual display for enhanced feedback.

The primary objective is to improve color recognition skills while also focusing on eye-muscle coordination and scanning abilities for individuals with visual impairments. Users will select colors using a color wheel and tap corresponding rings with the chosen color. For colorblind individuals we aim to improve their color recognition skills, whereas for visually impaired people, we aim to improve their gaze and movement around the table by performing interaction on the interactive table.

Throughout the design process, we focused on user-centered approaches, conducting research to understand the specific needs and preferences of individuals with visual impairments. This research guided the development of the system, ensuring that it addressed the unique challenges faced by this user group.

In this report, we will provide a detailed overview of the project, including the design process, implementation strategies, evaluation methodology, and the obtained results. We will also discuss potential constraints and difficulties that were encountered during the project, highlighting their impact on the final outcome.

By developing this interactive and engaging system, we aim to enhance color recognition skills and improve eye-muscle coordination for individuals with visual impairments and color blindness.

2. Project Description

As we mentioned earlier the goal of this project is to design and implement a multimodal system for color recognition targeting individuals with visual impairments and color blindness. The system aims to enhance their color recognition skills, improve eye-muscle coordination, and provide an engaging learning experience.

The multimodality of the game means that it uses several different modalities, such as speech and tangible objects, to provide a more complete and effective gaming experience.

In our case, these are speech and tangible modalities, which are used in the application to increase learning efficiency and ease of use. For example, the application can use sound and visual output to indicate the correct answer.

To achieve this goal, the project utilized an interactive table, which served as the basis for the game, and developed tactile objects that the user can touch on the table. The project revolves around the development of an interactive table-based system that incorporates two pucks and a microphone for user interaction. The interactive table serves as the central component which utilizes multiple modalities, such as tangible interactions with it and speech, to create a comprehensive learning environment. Users can select colors using a color wheel and tap corresponding rings with the chosen color. The system provides immediate auditory feedback on the correctness of selections. Also the project incorporates gamification elements to make the learning process enjoyable and motivating for the users. Together, these elements allowed users to play and interact with the game using different modalities, providing a wider range of options and ways of perceiving information, and improving the overall gaming experience for users with limited visual abilities.

The implementation of the system involved the utilization of hardware and software tools. The interactive table has sensors for object detection and touch recognition. It was connected with audio capabilities for speech prompts and feedback.

For the software aspect, Unity, a popular game development platform, was utilized to develop the software system. Unity allowed for the creation of interactive and immersive experiences, integrating the hardware components seamlessly. The software implementation included functionalities for object detection, touch recognition, visual feedback, and audio capabilities for speech prompts and feedback.

The project included a comprehensive evaluation phase to assess the effectiveness of the developed multimodal system. The evaluation involved measuring participants' color recognition accuracy, speed, and collecting subjective feedback on user satisfaction and engagement.

3. Design Process

a. Idea (Gamification, Version 1&2)

The game design and development process included a wide range of stages, starting from user needs research and game requirements determination, to testing and refining the final version. Each stage was essential in developing a multimodal game that would meet user needs and provide a more complete and effective gaming experience for users with visual impairments.

The design process of the project began with the generation of the initial idea, which centered around the concept of gamification and the development of two versions, Version A and Version B.

The idea of gamification aimed to create an engaging and motivating learning experience for the users. By incorporating game elements and interactive gameplay, the project sought to enhance user participation and enjoyment while improving color recognition skills.

Version A and Version B were devised as two different iterations of the system, each with its own set of mechanics and approaches to color recognition. These versions were designed to test and compare their effectiveness in achieving the desired learning outcomes.

The design phase involved brainstorming sessions, discussions, and conceptualization to outline the key features, functionalities, and visual aesthetics of both versions. Considerations were made to ensure the system's compatibility with the interactive table, two pucks, and microphone setup.

The design process also involved creating detailed design documentation, including user interface mockups, system architecture diagrams, and interaction flowcharts.

These documents served as a reference and guide for the implementation phase, ensuring a clear and systematic approach to the development of the multimodal system.

By focusing on gamification and designing two versions of the system, the design process aimed to create an interactive and inclusive learning environment that would effectively enhance color recognition skills for visually impaired and colorblind individuals.

The focus of this game on color blindedness, gaze and movement around the table and target the Visually impaired people for eye-muscle connection and scanning the environment.

We use two tangible objects(Puck A and Puck C) which are used to interact with our interactive table along with voice recognition. The Player will be required to select a color on the color wheel and later select the rings on the screen of the same color and confirm the selection using voice by saying the color of the rings.In order to complete a level, the player needs to remove all the rings by selecting and saying the correct color. To make the game engaging, we have added Stars to control the number of errors that can be made to pass to the next level. If the player loses 3 stars , they have one last chance before losing the game. We begin the game with 3 colors and add levels by increasing the number of colors to train on. We also decided not to incorporate a timer in the game as the goal is to train visual impairment people in their rehabilitation. By adding a timer, it would only stress them out and make them underperform.

Puck A is placed in the default position, and Puck C on a color wheel. Puck C can be rotated to select the color on the wheel. We provide visual aids to show the color selected while rotating the wheel. Once the color is selected using Speech we confirm the color. If the color selected on the wheel and the color spoken are correct, using Sonification, we confirm the selection, if its incorrect, the player loses a star. Once the color is selected on the wheel, the player can begin selecting the rings of the same color.

We perform our tests on 2 versions. In Version 1, the player selects each ring individually and confirms the selection by confirming the color. In Version 2, the play can select multiple rings of the same color and then confirm the color after selecting all the rings of the same color.

To design this game we used the CASE and CARE design space. Following were the techniques used to design our application.

b. CASE

In our design, we have implemented the CASE(Synergistic and Alternate Communication Types) model. The CASE model combines various communication types and modalities to create a comprehensive systems. By leveraging multiple communication types like, tangible objects and speech we provide the user with richer and immersive interaction. The use of multimodal communication opens new user interaction and engagement. For example, while selecting the color on the colorwheel using a tangible object and specifying the color via speech, provides confirmation and assurance for visually impaired people. Also, by providing auditory feedback with sonification, along with visual cues, increases players usability.

In our application, we have two main features, which requires two modalities. The two aspects are selection of color on the color wheel and selection of rings. With regards to the CASE model, we use Alternate and Synegistic in our game as shown in Fig 1.

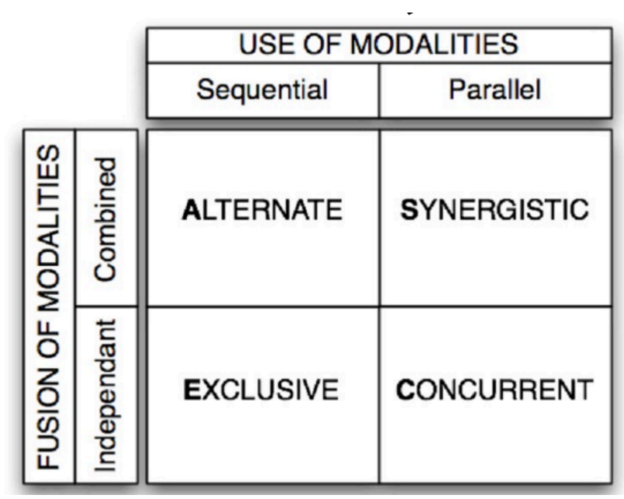

Fig 1: CASE MODEL

i. Version 1

<u>Selection of color on the colorwheel:</u>

The player needs to rotate and select the color with Puck C and then he confirms his selection by saying the color selected, by which we use the **Alternate** model, which is a combined fusion of the modalities and sequential use of the modalities.

<u>Selection of rings</u> :

The player uses Puck A to select individually each ring with the same color as setted on the color wheel and say the color. As the player needs to select the ring and say the color at the same time, we use a **Synergistic** model. The modalities are used in parallel.

ii. Version 2

<u>Selection of color on the color wheel:</u>
Similar to version 1, **Alternate** communication style is used for selecting the color on the color wheel, as the modalities are used in sequence.

<u>Selection of rings</u>:
The player uses Puck A to select multiple rings of the same color as setted on the color wheel and then after selecting all the rings, say the color. As this requires selection of the rings and then using speech to confirm its an **Alternate** model.

c. CARE

In our application, we use 2 properties of the CARE model to characterize and assess usability and fusion in our multimodal interaction application. They are complementarity and Redundancy.

Complementarity is the property where two modalities are needed to reach a given state. This property cannot be reached by using only a single modality.
For example, in order to remove a color ring from the screen, the player needs to select the ring using the tangible object as well as say the color of the ring.
Redundancy is the property, where using only one modality, a given state can be reached. For example, in selecting the color on the color wheel, by only using the tangible object we can select the color, but we add the fact of saying the color to confirm the selection.

Following are the properties used in our design:

i. Version 1

<u>Selection of color on the color wheel:</u>
We use the **Redundancy** property. This is because, while, rotating alone on the color wheel will be enough to confirm the selection of the color.
<u>Selection of rings</u> :
Selecting the ring individually using Puck A, and then saying the color, we are checking if the player has confirmed the selection of the ring thus making it **Complementarity** property.

ii. Version 2

<u>Selection of color on the color wheel:</u>
Similar to version 1, **Redundancy** property is used to select the color.
<u>Selection of rings</u> :
The modalities are **complementarity** to make sure multiple the selection of the rings are done by the Puck A and confirming with Voice.

d. Software architecture

The software architecture of the multimodal system, developed using the Unity game development engine, aimed to provide a seamless and immersive experience for color recognition. The architecture followed a modular approach, with different components serving specific functionalities.

The system integrated the following components:

1. Interactive Table with Visual Display: The interactive table, equipped with sensors for object detection and touch recognition, served as the primary interface for user interaction. It featured a built-in visual display that provided enhanced feedback, including highlighting the selected color and indicating correct or incorrect selections.

2. Pucks: The system utilized two pucks, puck A and puck C. Puck A enabled ring selection and tapping, while puck C functioned as a color wheel for color selection and rotation.

3. Microphone Integration: The microphone was crucial for speech recognition, enabling users to verbally select colors and confirm their choices. It facilitated speech input for both puck selection and rotation.

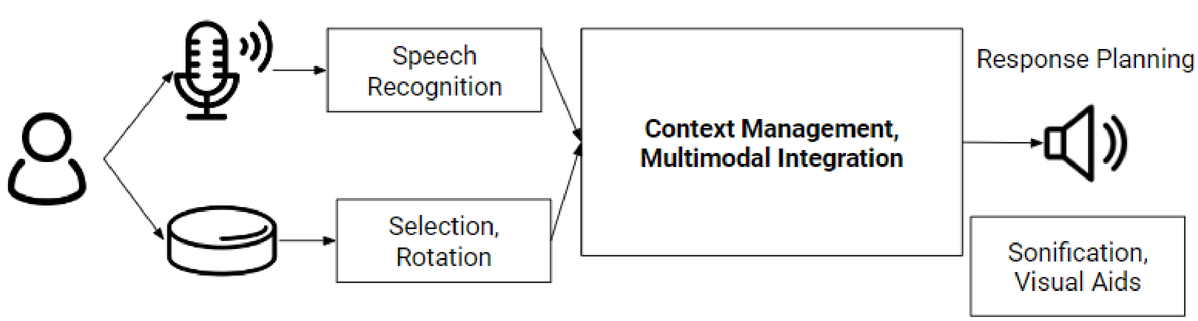

Fig 2: Software Architecture

In our project, the input modalities include the use of a microphone for speech recognition and tangible objects for selection and rotation. Following this, multimodal integration takes place using an interactive table, which serves as the primary interface for user interaction. It combines the data obtained from different modalities and provides users with feedback using output modalities such as sound and visual aids. This allows for a more comprehensive and continuous user-system interaction, enhancing the color recognition experience.

4. Implementation

To implement the project, a combination of hardware and software tools was utilized. The interactive table served as the central component, incorporating sensors for object detection and touch recognition. It was connected to a screen for visual output and integrated with audio output for feedback. The interactive table used was Ideum Platform II. This table came along with three tangible objects.

a. Tools

To build our application, we utilized Unity, a versatile game development platform, along with a tangible engine plug-in.

By using Unity as our development platform, we benefited from its extensive library of pre-built scripts, visual effects, and asset store resources. This allowed us to accelerate our development process and focus more on the core functionalities of our application rather than spending excessive time on low-level programming tasks.

Additionally, we incorporated a tangible engine plug-in into Unity. A tangible engine plug-in enables the integration of tangible user interfaces (TUIs) into the Unity environment. TUIs involve physical objects that users can manipulate to interact with the digital content. These tangible objects were placed on the interactive table, and the table's sensors detected the user's interaction with the objects. By integrating this tangible engine plug-in, we were able to create a more immersive and interactive experience for our users.

b. Conception

To build our application, we followed a series of steps after installing the tangible engine plug-in in Unity. One of the key features we utilized during development was the

Simulator mode, which allowed us to test and refine our application within the Unity environment before deploying it on the interactive table.

The simulator enabled us to interact with the application using virtual representations of the tangible objects. We added game objects such as a color wheel, initial positions, and rings, which needed to be interacted with the tangible object and speech.

To access and manipulate these game objects using the tangible objects, we implemented scripts that facilitated interaction between the two. These scripts allowed us to track the movements and gestures of the tangible objects and translate them into corresponding actions within the Unity application. For example, when the tangible object representing the color wheel was rotated, we wrote additional scripts to highlight the chosen color, making it easily detectable and providing visual feedback to the user.

Another important aspect of our application was speech recognition. To incorporate this feature, we had to enable microphone access and add Windows platform plugins within the Unity environment. This allowed us to integrate speech recognition capabilities into our game, which is the second modality of our application.

To combine the interaction modalities of tangible objects and speech, we developed a script that facilitated the integration of these two different input methods. This script enabled the selection of colors using tangible objects and confirmed the selection with voice commands. By writing this script, we achieved a seamless and intuitive user experience that combined the physical interaction with the tangible objects and the convenience of voice commands.

Utilizing these techniques, we successfully developed a captivating application that flawlessly merged tangible objects, speech recognition, and interactive gameplay into the Unity environment

c. Difficulties

Throughout the implementation phase, we encountered several challenges that required careful problem-solving and innovative approaches. One prominent difficulty involved ensuring accurate and reliable object recognition on the interactive table. Precise calibration of the sensors was crucial to detect and differentiate between the tangible objects effectively. In particular, we faced occasional mismatches between the highlighted color on the color wheel and the corresponding pointer, requiring us to rotate the tangible object to align with the selected color accurately.

Designing the tangible objects posed another challenge as we aimed to strike a balance between aesthetic appeal and functional usability. We had to consider both the visual attractiveness of the objects and their practicality in terms of interaction and manipulation.

Speech recognition presented its own set of challenges. We encountered issues with certain words not being recognized correctly, such as the color 'Red' which required us to substitute it with 'Burgundy' for accurate recognition. To ensure optimal performance, we utilized professional microphones with noise filters and repeated words in the presence of loud sounds. Additionally, accents proved to be a challenge, as the system needed to accommodate diverse speech patterns.

The installation process of the tangible-engine plugin proved to be less straightforward than anticipated. Insufficient documentation and lack of comprehensive examples made it challenging to reference and implement effectively. We had to create different profiles for the tangible objects to establish proper links with the 'Colorwheel' and 'Initial Position' components. Addressing lag issues during the rotation of the tangible object on the color wheel required a workaround of lifting and re-placing the object to mitigate the delays.

Establishing the connections between the game objects (such as Colorwheel, Initial Position, and Rings) and the tangible object also posed complexities. We had to carefully consider factors like the radius of the tangible object and the game object to ensure successful manipulation and interaction between them.

5. Evaluation

The main aim of our application is to enhance color recognition skills for people with visual impairments by using a multimodal system. Therefore, we want to evaluate our system.

Quantitative Evaluation:

a.Hypothesis

“Selecting specific color rings of the same setted color by taping and confirming each of them is slower but more precise than selecting multiple rings and confirming the group.”

b.Condition

To test our hypothesis statistically, we split it into two sub-hypotheses:

1. **Time :** Selecting individual rings will result in a lesser significant time than selecting multiple rings
2. **Errors** : Selecting individual rings will results in a greater significant errors

number than selecting multiple rings

The independent variable(IV) in this study is to “selecting color rings”. This IV has two sub conditions: (1) Individual selection of the rings and (2) Multiple selection of the rings.
Our dependent variables (DV) are : (1) Faster: time taken by the user to complete the level by removing all the rings and (2) Difficulty: the number of errors made while finishing the level.

We aim for a within-group experiment design and start with different versions for the participants to avoid any bias or memorization of the game.

c. Protocol

The steps for the evaluation process with the participants are the following:

- We start by individually talking to the participants
- We describe the aim of the game
- Provide explanation of version 1
- Start the game for them, by selecting ‘Play’
- We time the participants on the how long it takes to complete a level
- We collect some feedback and note the errors made
- Next, we provide them with the explanation to Version 2
- Start the game for them, by selecting ‘Play’
- Time the participants to complete a level
- We collect feedback and note the errors
- After collection of feedback ask their preference of their versions.

Here are the guidelines provided to the participants:
“Hello, welcome to Tap!Tap! A rehabilitation game for visually impaired people for helping with color blindness and eye muscle movement. We train them to differentiate 8 colors. Namely, ‘ burgundy, yellow, orange, blue, green, pink, purple and grey..’ The aim of the game is to remove all the rings on the screen by identifying their corresponding colors by voice. We can begin the game by selecting the play button, then place Puck A in the initial position, the white circle and Puck C on the color wheel. By rotating puck C select the color and say the color. Then using Puck A, we can select the rings. To change the color setted on the color wheel, Puck A needs to be brought back to initial position. We have 2 versions. In Version 1, Puck A is used to select a ring and say the color individually, "Do you have any question before we start the game?Will start the timer when you are ready” [Participants play the first version].

“In the next version, after setting the color on the wheel with Puck C, multiple rings can be selected using Puck A and then confirm the color by saying their color. Any questions? Will start the timer when you are ready”

d.Evaluation

In our controlled evaluation, we conducted a "within-group" design to assess our hypothesis. We selected a total of 8 users for the evaluation. Both male and female users were included in the study to ensure a balanced representation.

The evaluation measured participants' accuracy and speed in identifying colors. It also included subjective feedback through questionnaires to assess user satisfaction and engagement with the multimodal system.

By employing this controlled evaluation approach, we aimed to gather quantitative and qualitative data from the users, allowing us to assess the performance and usability of our multimodal game for rehabilitation purposes for the visually impaired.

6. <u>Results</u>

Following are the results obtained from our evaluation. Below is the table with the amount of time taken by the participants to complete the game as well as the errors made. We evaluated the participants on both the versions.

Participants	Version 1: Synergistic		Version 2: Alternate		Started Synergistic
	Time	Error	Time	Error	
1	0:02:15	2	0:01:11	0	True
2	0:01:09	1	0:00:42	1	False
3	0:01:35	2	0:00:41	3	True
4	0:00:55	0	0:00:48	1	False
5	0:01:30	3	0:00:34	0	True
6	0:01:54	5	0:01:03	1	False
7	0:01:38	0	0:01:10	1	True
8	0:01:39	0	0:02:09	1	False

Table 1: Time and Error for Version 1 and 2

Global	V1: Time	V1:Error	V2: Time	V2: Error
avg	0:01:34	1.625	0:01:02	1
min	0:00:55	0	0:00:34	0
max	0:02:15	5	0:02:09	3
std	0.0002863226023	1.767766953	0.0003514812646	0.9258200998

Table 2: Statistics on the global data

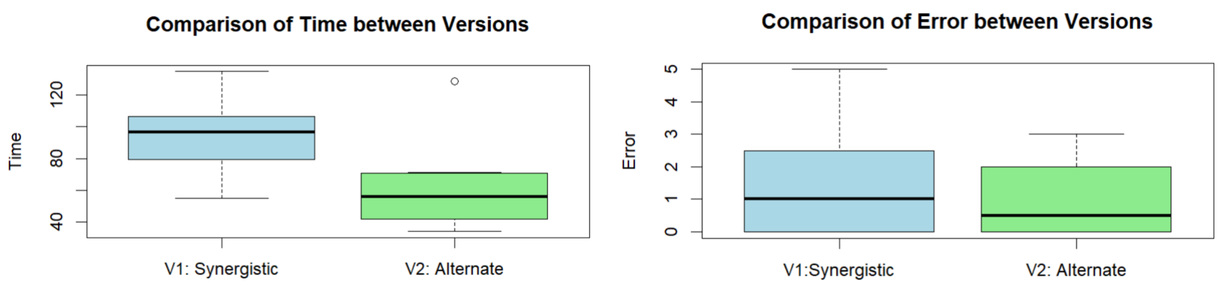

Fig 3: BoxPlot on Time and Error comparison between Versions

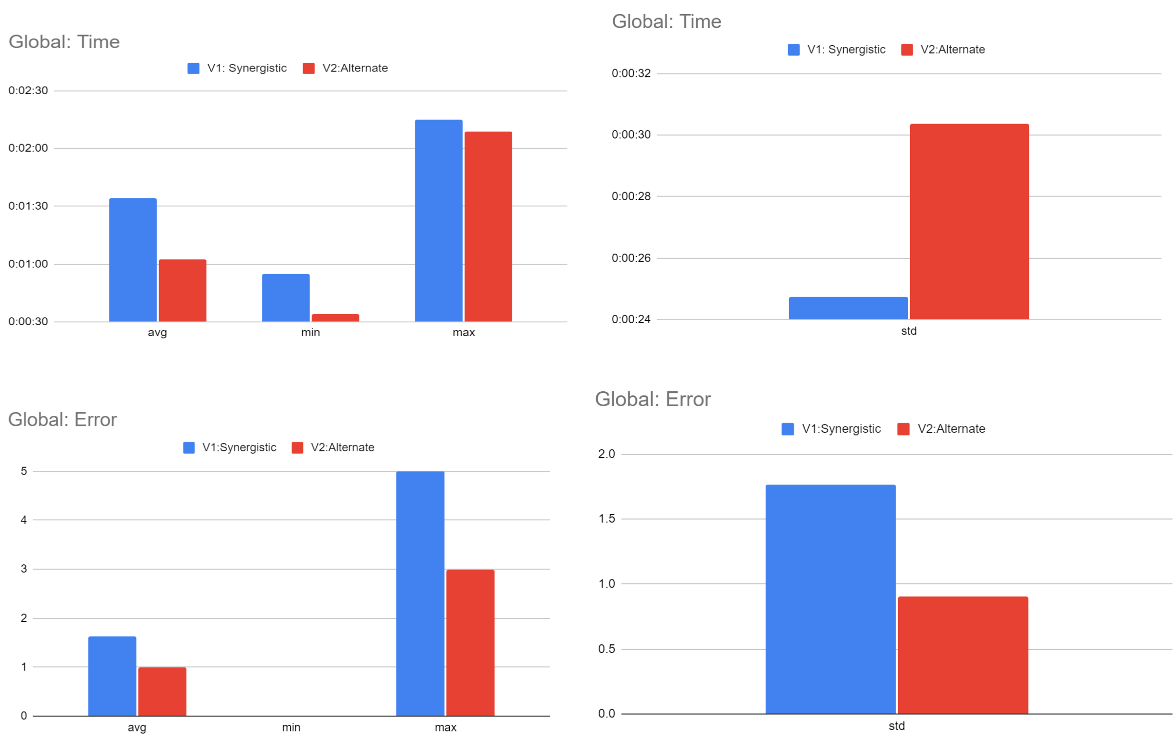

Fig 4: Min, Max, Average and Standard deviation for V1:Synergistic & V2: Alternate

In order to know if our observation has statistical significance, we performed a paired T-test on the collected data. But since the number of participants are only 8, we cannot say our results are statistically significant. To perform paired t-test we cannot remove the 4 outliers, so we decided to keep them to do our paired t-test.

Results from R: For global values

Welch Two Sample t-test

data: time_v1 and time_v2 **t = 2.3198 secs, df = 13.45, p-value = 0.03666** alternative hypothesis: true difference in means is not equal to 0 95 percent confidence interval: 2.309073 secs 61.940927 secs sample estimates: Time differences in secs mean of x mean of y 94.375 62.250	data: error_v1 and error_v2 **t = 0.88586, df = 10.571, p-value = 0.3954** alternative hypothesis: true difference in means is not equal to 0 95 percent confidence interval: -0.9355711 2.1855711 sample estimates: mean of x mean of y 1.625 1.000

Table 4. shows the aggregated qualitative feedback provided by the participants.
Description of the columns:

1. Participants : serialize the participants to anonymize the results of the participants
2. Feedback v1 : After playing the Version 1, we ask the feedback from the participants
3. Comment : Additional comments they had
4. Feedback v2 : After playing Version 2, we ask the feedback
5. Comment : Additional comments from their point
6. Start with V1 : Here we check which version the participant started with
7. Preference : we asked their preference to the version played and why.

#	feedbacks v1: synergistic	comment	feedbacks v2: Alternate	Comment	start with v1	preference	pref
1	"not sure of the name of the colors, according to the palette (forget to have to say the	went always initial pos that triggers color	nice the seen to say that it is selected, sound for the selection confirmation --> on the circle (audio or visual	voluntarily missed for the test	TRU E	version 2, funny to select multiple times, more interactive and	2

	range), color wheel not align at first don't know when you're on it if the ring color is gone or not, otherwise funny"	selection, then directly tapped , sound was not there	feedback), when you end on false, trigger false, when you end on true when there is false, sound true --> audio feedback on last selection but correct visual			dynamic than talk every time, maybe unsure that selection is take in account	
2	"frustrating, can't go as fast as you would like, use only the left hand when theoretically the right hand is free. more logic in the frame than in version 2, more intuitive Focus, because of the large screen, not intuitive scanning, no complete representation, even for non-VIP, understands the training task. More colors for both versions would be interesting"	wasn't going to say color before tap	little interesting handling, not intuitive to keep puck A before confirming selection (thought I had to always return to the initial position), redundancy on say 2x, type directly	reset initial position before confirming the choice	FALSE	version 1, weird and fun to say multiple times and every time. force to take your time. concept tux of math comment, educational game, same feeling	1
3	doesn't understand the accent or program, tactile is good, rotation and placement are working fine, vocal and confirmation are a pain	-	moment to understand the systematic, more flexibility, need to follow logic is too restrictive	-	TRUE	alternate (v2), faster, especially because voice problem	2
4	wheel is fun but sometimes a pain	-	even if put back initial position when selected multiple rings, not set	-	FALSE	alternate (v2), faster and easier because	2

			color (of indicate lift it up)			less voice command	
5	easy to use once you understand	-	well done	-	TRUE	don't see difference but 2nd because more efficient	2
6	easy , fun, interesting, felt real, easy to manage, more colors, difficulty saying the color	-	fun, easy, interesting	-	FALSE	v2, less repeating colors	2
7	have own color panel, if follow instructions, it's easy	-	not too hard, didn't fully understand the mechanism (put the puck at initial position after confirmation to the selection color mode)	error not said color, forgot to select all the rings in one but did it twice	TRUE	v1, more logical, as the other one isn't consistent in saying the colors	1
8	all possible colors should be usable, app not bad	reflection on table might impact the color	color names could have been easier, putting the puck back at the initial position instead of confirming and then put it back would have been more logical, too many redundant confirmation, should have better sound recognition	went back to color selection mode before confirming his choice	FALSE	v2 : would have been even better if no need to have the final confirmation before changing the color, reason not 1st is because of poor voice recognition (need to talk too loud and be too careful with pronunciation)	2

Table 3 : Qualitative feedback

7. Discussions

Evaluation of Quantitative analysis

We have aggregated our recorded data and present them in Table 1. We perform statistical analysis on the data and present them in Table 2. We can observe that the time taken to complete the level in Version 1 is higher meaning it is slower. With a within-group experiment design, preferred for our small sample size, one IV (selecting a ring), two conditions (Individual or multiple ring selection), and two dependent variables (DV) (time, error), we perform a paired-samples t-test to test our hypotheses as mentioned in Section 5.

We perform the paired-samples t-tests for Time (p-value = 0.03666) and Error (p=0.3954), with null hypothesis of the true mean difference between the paired samples being zero. At the 5% significance threshold, we can observe that the null hypothesis can be rejected, and say that there is difference in the mean. Hence our sub Hypothesis on Time as mentioned in Section 5, can be rejected, and we observe that Version 1(synergistic) is slower compared to Version 2(altenate). The reason why version 1 is slower, is that participants need to individually select the ring with the tangible object and then say the color for each of them. As the speech recognition does not recognize the commands well, it takes more time making Version 1 slower to Version 2.

For the sub hypothesis on error, as the p-value is higher that 5% threshold, it cannot be rejected. Thus, there is a higher error rate while selecting individual rings as compared to selecting multiple rings. Reasons errors can not be rejected as it should be because we did not take into account the voice recognition errors. This is due to the quality of recognition, the intensity with which participants had to say the colors and the time taken to trigger which all impacted the application.

Our within group controlled experiment showed us insights into the designed application. But as we had only 8 participants to evaluate, we cannot say that they are statistically significant results.

From our analysis we can conclude that, Version 1 is more intuitive but less preferred as it takes longer to execute because the voice recognition dictionary takes a while to recognize and needs to be precise. Whereas Version 2 bypasses that by confirming less but logic or procedure is harder to grasp as the voice modality isn't consistent.

Qualitative Evaluation Analysis:

Based on the feedback obtained from the participants as shown in Table 4, here are some of the key findings.

1. **Version 2 more preferred** : Participants preferred version 2, mainly as speech recognition is prone to misinterpret the commands
2. **Sonification** : Participants liked the feedback obtained as sounds when the color selection is made. Participants also wanted more feedback, as to know if Puck A has been detected or not.
3. **Color identification** : Participants found it hard to specify the color
4. **Fun and interesting** : To use the tangible object and speech
5. **Ease of use** : Participants provided positive feedback on ease of use when they were informed on how to play.
6. **Lag on color wheel** : Some participants incurred lag when setting color on the color wheel and found it frustrating.
7. **Initial Position** : Participants found placing Puck A in the initial position to set the color confusing.
8. **Speech Recognition** : Participants had to be loud and clear while providing the colors.

Overall, the qualitative evaluation provided insights into users perceptions and opinions regarding the use of tangible objects and speech modalities.

8. Conclusion

In conclusion, we have developed a multimodal application aimed at enhancing color recognition skills for individuals with visual impairments. Through the integration of tangible objects and speech recognition, our application provides a seamless and immersive user experience.

The implementation of the CARE (Complementary, Redundant, and Equivalent) model in our design allowed us to leverage multiple communication types and modalities, providing a comprehensive and engaging interaction for users. The use of tangible objects for selection and rotation, combined with speech recognition for confirmation, enabled users to interact with the application in a more intuitive and inclusive manner.

The software architecture of our application followed a modular approach, integrating components such as the interactive table, pucks, and microphone integration. Unity and the tangible engine plug-in were instrumental in developing and integrating the different modalities seamlessly.

During the implementation phase, we encountered challenges related to object recognition, designing tangible objects, and speech recognition accuracy. However, through innovative problem-solving and iterative development, we were able to overcome these difficulties and create a functional and captivating application.

For future enhancements, we recommend incorporating more sonification to provide auditory feedback, adding a tutorial to guide users in color recognition, and introducing gamification elements with different levels to make the learning experience more enjoyable and challenging.

Overall, our multimodal application shows promise in assisting individuals with visual impairments in improving their color recognition skills. Further evaluation and user testing will help refine and optimize the application to better meet the specific needs and preferences of the target users.