Testing Purposes and Objectives

The purpose of the tests is to obtain feedback from users to further improve the current prototype design. Goals include finding bugs, gathering opinions about the features and activities, and collecting data on how helpful the features and activities are so that changes can be made to them. Furthermore, the user will be given the opportunity to make suggestions on features and activities to be added in the future.

Objectives:

1. Complete the Ease of Use/Accessibility Test
2. Complete the Language Immersion & Activities Test

Testing Procedure

Ease of Use/Accessibility Test

For the first test the users will be split into two separate groups:

Group A (Able to ask questions and be given assistance) will be given the VR headset and will first explore the game menu which will include learning the game controls, watching video demonstrations of the activity, and testing the functionality of the menu. Then they will have time to explore the open world environment on their own for 15 minutes with the goal of interacting with as many game objects and activities as possible. They will then be told a set of tasks to perform and then they are required to return to the game menu.

Group B will be given the same instructions, but they will not be able to ask questions and will do all the tasks without any assistance.

Language Immersion & Activities Test

Test Hypothesis: Using Foreign Language Immersion Virtual Reality (FLIVR) will lead to faster language vocabulary learning and longer language retention than traditional language learning methods.

The purpose of this test is to compare how effective FLIVR is at helping the user memorize vocabulary compared to traditional methods. Because this test involves language retention, it will take place over the course of three days. The participants in Group A will spend 30 minutes using FLIVR and group B will spend 30 minutes studying flashcards of the same game objects that Group A would encounter will using FLIVR. Each day after the thirty minutes is up, the participants will take a test to track their progress.

Group A will be using FLIVR with the purpose of learning language as close to how it naturally happens. FLIVR is meant to imitate the natural way that people learn languages, so group A will not be expected to repeatedly identify objects unless they want/feel the need to do so. The activities will incorporate the words they have discovered so far and will use a spaced repetition algorithm to determine how often the user will see vocabulary words in the activity.

Group B will be given use of a web program that features a spaced repetition algorithm. Compared to regular physical flashcards, ones that use a spaced repetition algorithm show much higher retention and faster initial learning speeds than traditional flashcards.

Place and Time

The Ease of Use/Accessibility Test will take place in a public setting, like a classroom or study room, and should take no longer than 30 minutes. The Language Immersion and Activities Test will take place in their home (or wherever they find comfortable), and group A will be given a VR headset to use in their own desired time. Both groups will be instructed to only use FLIVR or study flash cards.

Participants

There will be 10 participants in this experiment with various skill sets and backgrounds. Although each participant has different skills, they must all have the following requirements:

1. Fluency in the English Language
2. The physical ability to operate a VR headset.
3. The willingness to participate and answer questions honestly.
4. The ability to share their opinions on the features the application offers and to make suggestions for new ones.

Experimental Design

The tests will focus mainly on two areas: tests of language retention and tests of application ease of use/accessibility. The primary goals of these tests are to obtain actionable results that will help improve the effectiveness and usability of the application.

Independent Variables - We will categorize test participants based on the following independent variables:

- Oculus Quest/VR experience.
- Spanish Language experience.

Dependent Variables - Performance measure will be based off the following:

- Scores achieved in the activities.
- Time taken to achieve the scores.
- Number of input errors when completing the multiple-choice activity.

Data Collection

The main method of data collection will be the questionnaire that is provided after the tests have concluded, as well as observations made after viewing the recordings of the tests. Additional data will be recorded using the JSON database. These methods will allow for quick generation of easy-to-understand graphs that display reaction times when using the application. However, this metric might have to be cut due to time constraints. If this system is cut, we will time participants manually to gather this data. Finally, at the end there will be a very short, open-

ended interview with each user after their test to get any feedback that they might have on the application.

Result Analysis

When analyzing the results, negative trends in the data will be studied. All data collected will be analyzed for actionable results that will improve the application in some way.

Questionnaire Questions:

Yes or No questions to judge participant's prior experience:

Have you used VR before?

Have you used an Oculus product before?

Have you ever studied foreign languages before?

Users will rate the next set of questions from 1 to 5 with 1 being Strongly Disagree and 5 being Strongly Agree:

The in-game instructions were easy to understand.

The main menu was easy to navigate.

The Heads-Up Display (HUD) was easy to understand.

The audio pronunciations of words are easy to understand.

Flower Activity was engaging.

Multiple Choice Activity was engaging.

The app mimics language immersion well.

The app helps the user learn their target language.

These Questions are short answer:

Did you experience any confusion while using the app, if so, when?

Were there any actions you wanted to take but could not?

Which interaction did you find the most difficult to complete?

What features do you think would be beneficial to add?