University of Glasgow | School of Computing Science

VIA VERITAS VITA

A WORLD TOP 100 UNIVERSITY

Visual Field Testing In Virtual Reality With Eye Tracking and Visual Feedback

By Alfie J. Macintosh (2557035m)

WORLD CHANGING GLASGOW

THE TIMES THE SUNDAY TIMES GOOD UNIVERSITY GUIDE 2024 SCOTTISH UNIVERSITY OF THE YEAR

Outline

- Introduction, Aims and Motivation
- Our Design
- Our Implementation
- User Evaluation and Data Collection
- Usability Results and Impact of Visual Feedback on Test Results
- Individual Case Study
- Conclusions

Participant (Top-Left) during evaluation – consented to photo. (Top-Right) Main Menu, (Bottom) Vive HTC Headset & Controllers

Introduction

- Visual Field Testing is a diagnostic test used to measure the full range of what a user can see (their visual field)
- Used to diagnose conditions several medical conditions
- Also used to track the progression of conditions
- History:
  - Hippocrates (5 Century B.C.)
  - Dr. Hans Goldmann (1940's)
  - Dr George Lynn and George Tate (1970's)
  - David and George L. Humphrey (1980's)

Motivation

- Accessibility and cost
- Patient Compliance
- Current Virtual Reality Technology
- Possibility to enhance the reliability of visual field testing

Aims

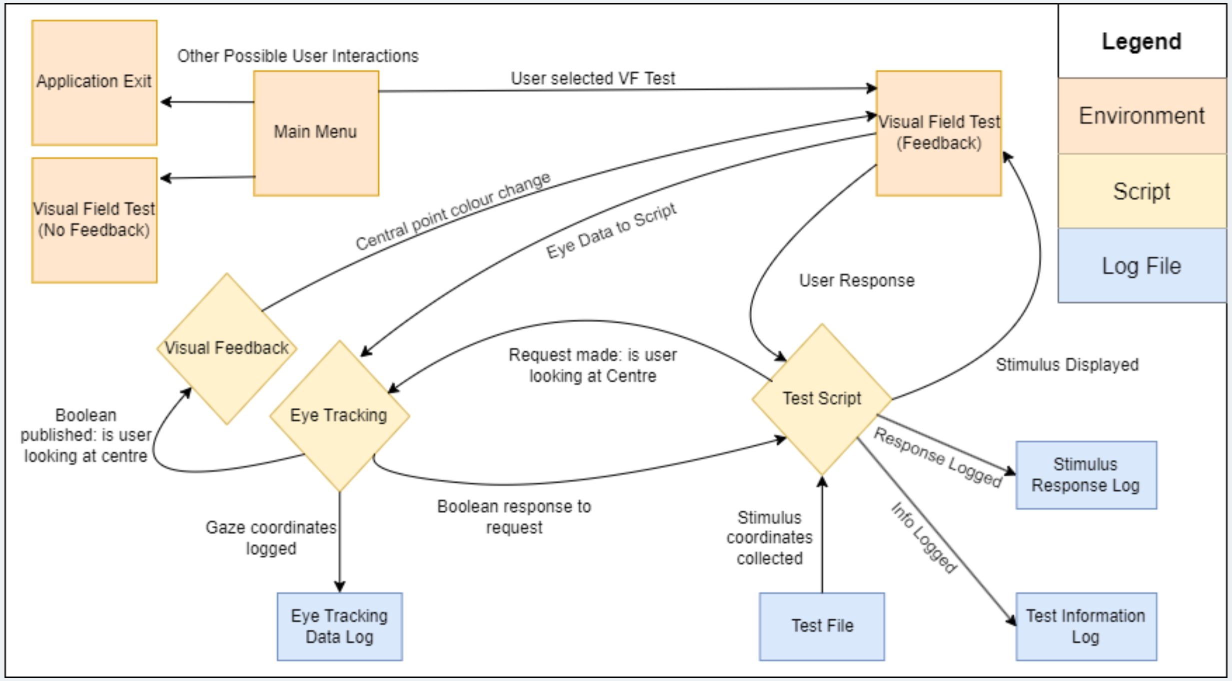
- Develop a visual field testing platform in virtual reality
  - With eye-tracking functionality
  - And real-time visual feedback to the user
- Enhance the reliability of visual field testing in VR measured by key metrics
  - While maintaining usability

Design

Overview

- Three environments
  - Main Menu
  - Test (with visual feedback)
  - Test (without visual feedback)

- Interactions
- Input/Output
- Dependencies
  - Test Script
  - Eye Tracking
  - Visual Feedback



Test Script

- Input
  - Test file
- Output
  - Results output
  - False positives
  - Test info

- Our pseudocode in Algorithm (1) outlines functionality of our test script. Furthermore, we can see the test algorithm we will implement

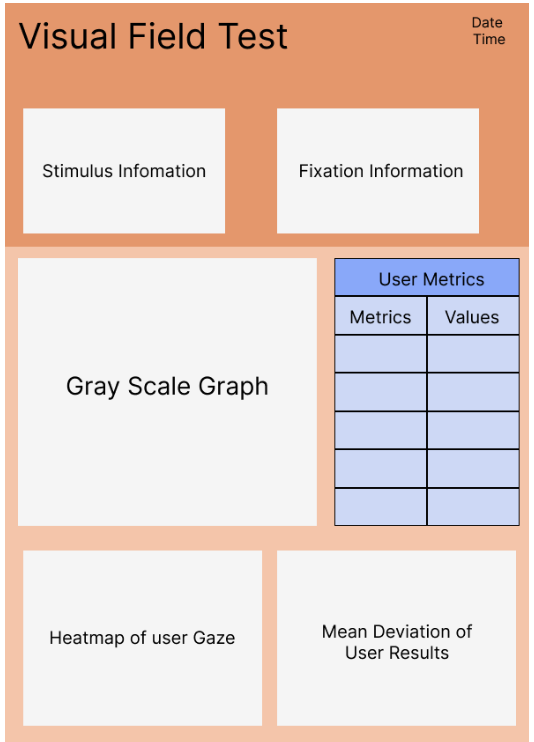
Finally, we need to process our test data...

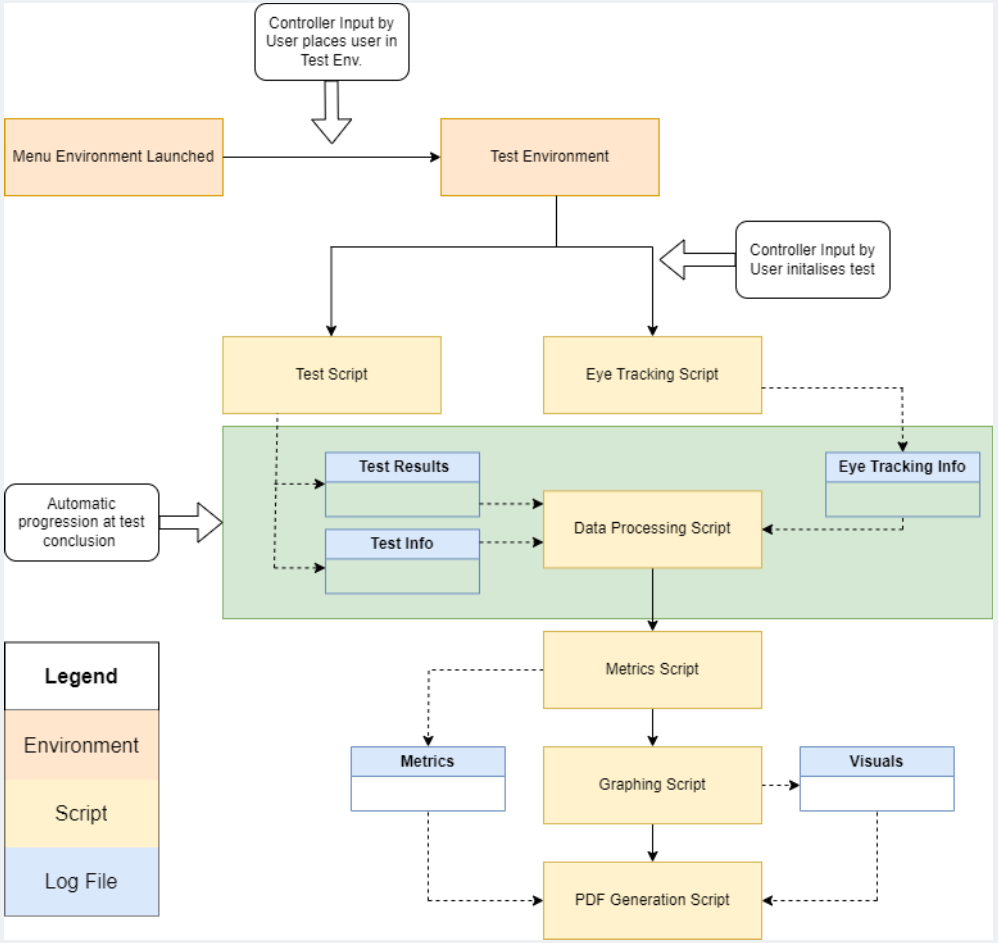
```
Input: Test File containing number of stimuli, response time, and stimuli positions
Output: Log files for results, false positives, and test info
/* Setup initial references and start test                                  */
Reference Canvas;
Reference Eye-Tracking;
Start Test and Record Start Time;
Setup File Paths for Logging Output;
/* Read Test File and parse parameters                                      */
Read Test File: Parse Number of Stimuli, Response Time, and Stimulus Positions;
Setup Writers for Results, False Positives, and Test Info;
/* Main loop for displaying stimuli and logging results                     */
foreach Stimulus do
    Display Stimulus at Current Position on Canvas;
    Wait for Trigger;
    if Triggered Within Response Time and Looking at Centre then
        Log True for Response;
        Log True for Looking at Centre;
        Log Reaction Time;
    else
        if Triggered Within Response Time and Not Looking at Centre then
            Log True for Response;
            Log False for Looking at Centre;
            Log Reaction Time;
        else
            Log Missed Response;
        end
    end
    Wait Random Interval Before Next Stimulus;
    if Trigger During Waiting then
        Log False Positive;
    end
end
/* Clean-up Logging                                                         */
Stop Eye-Tracking Script;
Log Results for Each Stimulus and False Positive;
Close Writers;
Close All Open Files on Application Quit or Disable;
```

Algorithm 1: Visual Field Testing Algorithm

Report Processing

- Raw -> Processed (Graphs)
- Raw + Processed -> Report





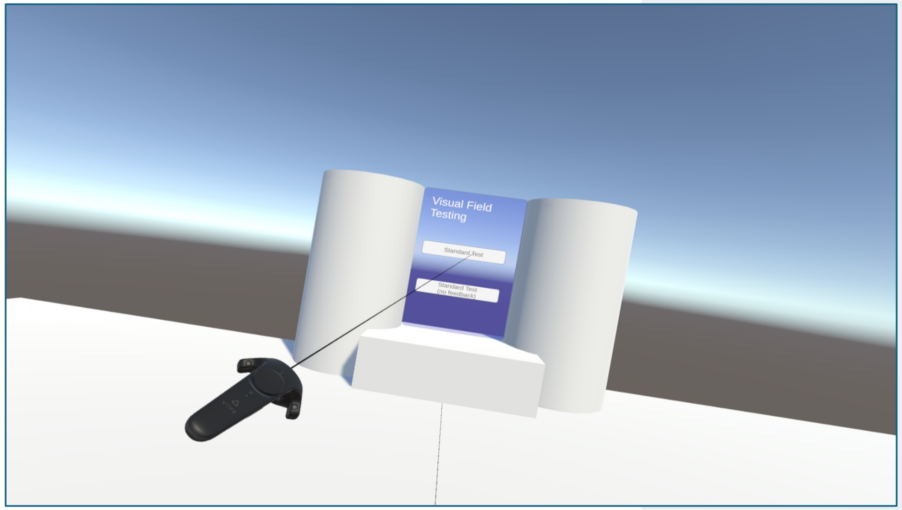
Implementation

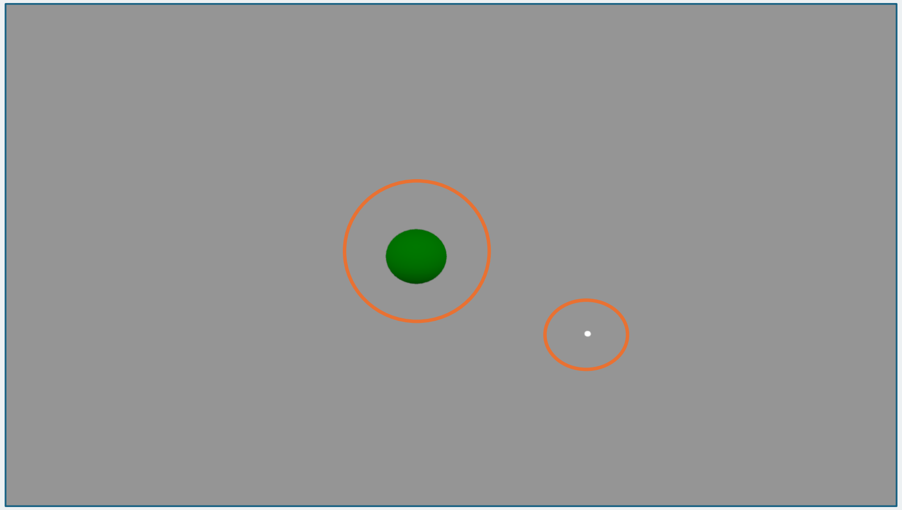
Hardware & Software

- Hardware, we used the following:
  - HTC Vive Pro Eye headset due to its eye-tracking capabilities
  - HTC Vive controllers 2x (however, either one would work individually)
  - Eye-Patches to isolate testing to one eye

- Software, we used the following:
  - Unity
    - SteamVR
    - SRAnipal
    - Unity (MonoBehaviour)
  - Unity scripting was done in c#
  - Graphing and reporting was done in Python

Environment





User is looking at centre

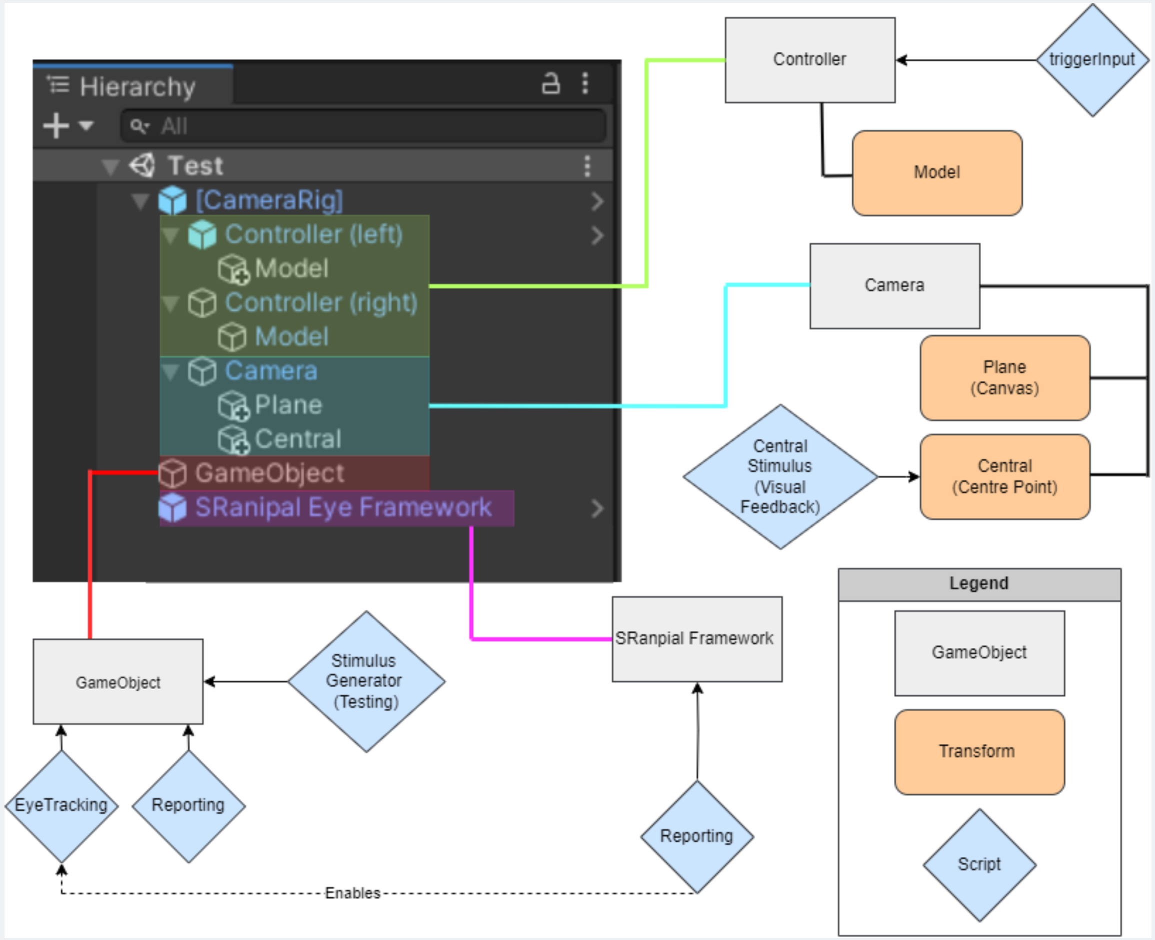
User is not looking at centre

Test is in ‘idle’ mode

Unity Program

Example Scene 'Test'

- Unity is hierarchical
- gameObjects function as containers for logic and models



Reporting

Participant 1 – Left Eye

Visual Test 2024-03-18 16:59:14

Test Duration: 362.91s

Stim Info

Type: Default White 0.05

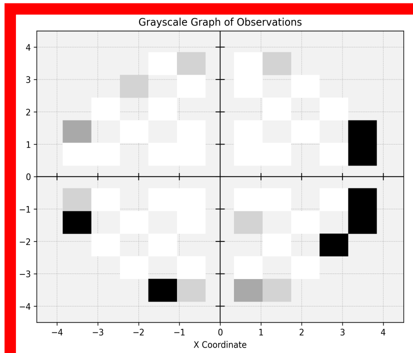
Canvas: Plane

Display Time: 80.0ms

Fixation Info

Fixation Monitor: Gaze

Fixation Target: Central



Metrics

False Positive Error Rate: 5.13%

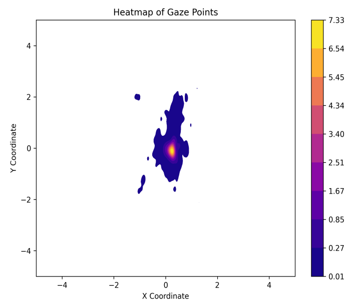
False Negative Error Rate: 6.45%

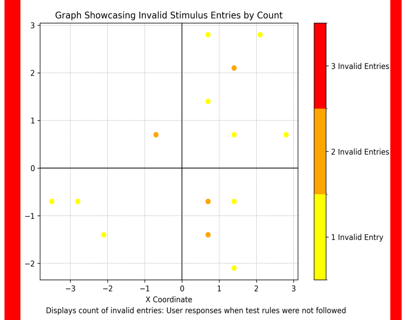
Average Reaction Time: 462.97 ms

Total Responses / Total Stimuli: 124/156

Valid Responses / Total Stimuli: 106/156

Time looked at Centre: 83.93%





Evaluation

All photographed participants gave their consent and were informed of its use in this presentation

Methodology

Hypothesis: Visual Feedback during testing will increase common validity metrics found in visual field testing such as false-positive, false-negative and fixation rates. Furthermore, the percentage of points registered that are 'valid' will be higher.

A valid point requires the user (1) registers it, (2) is looking at the centre. We further wanted our virtual reality system to be usable.

- Eight Participants in our user evaluation and one standalone study
- We conducted our evaluation on usability with all participants
- When comparing our two test conditions with and without visual feedback we excluded the results from our standalone study as the participant had previous experience with visual field testing
- Participants wore an eye patch to isolate testing to one eye per test
- They will use the same test condition throughout both tests

Usability Results - Cybersickness

- In phase 1 88.9% of participants had eyestrain symptoms. In phase 2, this decreases to 66.6%.
- General discomfort and drowsiness typically resided after the first phase of testing apart from in one participant who increased drowsiness to 'Somewhat'
- Symptoms of headaches were reported in two participants

Table 6.1: *Comparison of Eye Strain Frequency across both testing phases*

Eye-Strain Frequency	Not at all	A little	Somewhat	Often	Throughout
Phase 1 (First Eye)	11.1%	55.6%	11.1%	11.1%	11.1%
Phase 2 (Second Eye)	33.3%	11.1%	22.2%	22.2%	11.1%

Table 6.2: *Comparison of General Discomfort Frequency across both testing phases*

General Discomfort Frequency	Not at all	A little	Somewhat	Often	Throughout
Phase 1 (First Eye)	44.4%	-	44.4%	11.1%	-
Phase 2 (Second Eye)	55.6%	22.2%	22.2%	-	-

Table 6.3: *Comparison of Drowsiness Frequency across both testing phases*

Drowsiness Frequency	Not at all	A little	Somewhat	Often	Throughout
Phase 1 (First Eye)	55.6%	44.4%	-	-	-
Phase 2 (Second Eye)	77.8%	11.1%	11.1%	-	-

Usability Results – Equipment Comfortability

- In phase 1 most participants rated the VR headset as 'somewhat comfortable' (55.6%) followed by 'neither comfortable nor uncomfortable' (22.2%). A small percentage of participants found the VR 'somewhat uncomfortable' at (11.1\%) and 'very comfortable' (11.1%). In Phase 2 we see the distribution shift, with a higher percentage of participants rating it 'neither comfortable nor uncomfortable' and 'somewhat uncomfortable'
- The eye patch was initially rated less favourably than the headset with 77.8% of participants finding it uncomfortable in some capacity. The distribution shifts slightly in the second Phase with 44.4% of users finding it uncomfortable in some capacity, however, it is never rated 'very comfortable' and only 2 participants rate it as comfortable in some capacity.

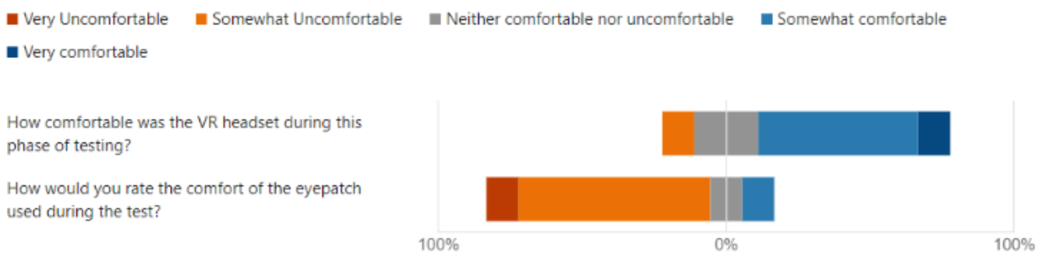


Figure 6.1: Equipment comfort evaluation for Phase 1

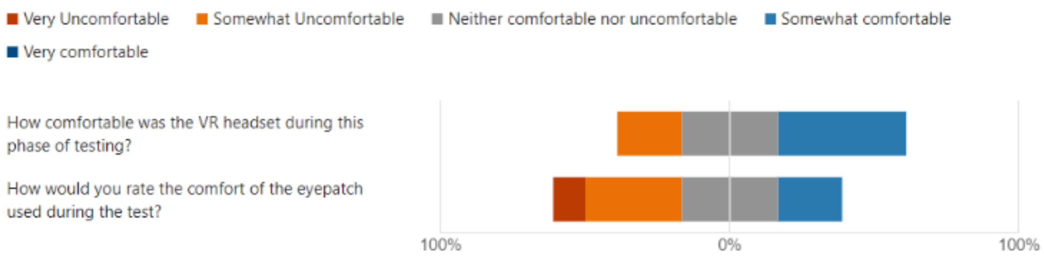


Figure 6.2: Equipment comfort evaluation for Phase 2

Usability Results – Testing Duration

- The mean test time was 454.925 seconds
  - lower bound of 362.91 seconds
  - upper bound of 731.96 seconds.

- 89.9% of participants found test duration reasonable with 11.1\% finding it unreasonable.

- Comments such as:
  - 'at the least minute the eye strain was a little too much. 5 minutes [of testing] at most?'
  - 'slightly too long. could do a shorter test?'
  - 'reasonable, but I would have preferred it to be faster as it started to get slightly uncomfortable'

- suggests although participants may have found the test reasonable there is a strong demand for slightly shorter test times, perhaps around the 5 minute mark, or in our case, around the lower bound of our test time.

Overall Usability

Table 6.4: How would the user rate their overall experience with visual field testing in VR

Rating	Very Poor	Poor	Average	Good	Very Good
User Responses	-	-	-	89%	11%

Potential Improvements:

- Quick Test Duration
- Removal of Eye Patch from methodology
- Habituation period to virtual reality

Statistical Analysis

T-Test:

H0: There is no significant difference in the mean percentage of valid responses between the condition with visual feedback and the condition without visual feedback

H1: The mean percentage of valid responses is significantly higher in the condition with visual feedback compared to the condition without visual feedback.

Not without limitations...

$Mean_{VF}$: 98.21034273
$Mean_{NVF}$: 91.91749228

$StandardDeviation_{VF}$: 2.599364617
$StandardDeviation_{NVF}$: 8.780507569

Given the results from our mean and standard deviation calculations across both test conditions and using the t-test formula we conclude that our **t-test equals: 1.943705608**. As degrees of freedom (df) is equal to: $df = n_1 + n_2 - 2$ we can use our degrees of freedom (df=6), our one-tailed t-test (α) and our significance level (p=0.05) to consult a T-Distribution table of critical values[1] to find the critical value.

Given the critical value is: 1.943, and our t-value was: 1.944 (3 s.f.)

Table 6.5: Mean values of validity metrics across both test conditions.

Test Condition	False-Positive (%)	False-Negative (%)	Time looked at centre (%)
Visual Feedback	5.05	13.65	98.34
No Visual Feedback	5.53	7.28	91.92

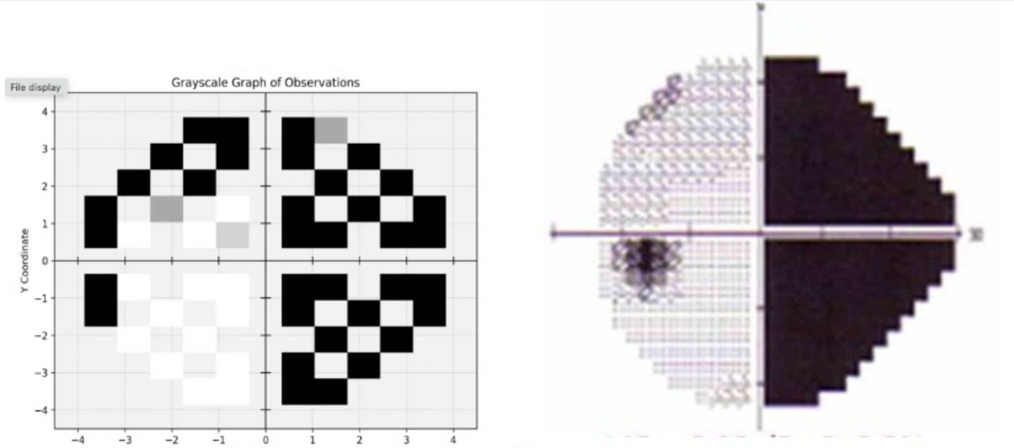
Limitations

- Small Sample Size (n=8, for statistical analysis)
- A very small difference between critical value and t-test value
- Small difference in false-positive rate (%)
- By changing eye patch after eye tracking calibration, eye tracking data could be incorrect

Individual Case study

- User with right-side homonymous hemianopia
- Metrics of:
  - Mean fixation rate: 94.6%
  - False-positive rate: 3.53
  - False-negative: 15.48%

- commented that there was a ’bleaching effect’ when looking at the central stimulus
  for too long. They suggested that the central stimulus change shape so this effect is mitigated



(a) Gray scale from standalone participants left eye.

(b) Sample gray scale from the left eye. Retrieved from Perez and Chokron (2014).

Conclusion

- This project achieved a visual field test within virtual reality, using eye tracking to provide visual feedback to users, when they weren't following testing conditions
- There are statistically significant results highlighting that visual feedback provides a greater valid point mean percentage than tests without visual feedback ($p = 0.05$)
- We outlined limitations of our evaluation including small sample size (n=8, 16 tests) and the flaw in our methodology when users are asked to remove the headset to put on their eyepatch
- Several users had eye strain. This appeared to resolve for most participants, but in some, it became more severe
- The eye patch was found to be uncomfortable by the majority of participants.

University of Glasgow | School of Computing Science

Thank you.

I would like to acknowledge:

- My supervisor, Paul Siebert for helping guide me throughout this project
- Fergus and Paulius for keeping the office an interesting place to be