Tracking Eye Behavior in Decision-Making Maze Problems

by

Jing Xian Wang

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Engineering in Computer Science and Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2019

Author . . . .
Department of Electrical Engineering and Computer Science
May 24, 2019

Certified by . . . .
Randall Davis
Professor
Thesis Supervisor

Accepted by . . . .
Katrina LaCurts
Chairman, Department Committee on Graduate Theses

Tracking Eye Behavior in Decision-Making Maze Problems

by

Jing Xian Wang

Submitted to the Department of Electrical Engineering and Computer Science
on May 24, 2019, in partial fulfillment of the
requirements for the degree of
Master of Engineering in Computer Science and Engineering

Abstract

Neurodegenerative diseases, such as Alzheimer's and Huntington's, can significantly impact a patient's ability to solve everyday problems. The ability to detect early signs of mental decline is crucial for determining whether someone might be at risk for these diseases. Eye behavior is often correlated to cognitive load, so examining the behavior of the eyes during decision-making tasks could provide us with insights on how individuals think when solving challenging problems. But eye tracking data is notoriously noisy, so developing methods for cleaning, parsing, and visualizing the data is important for understanding what it means. By conducting eye tracking studies on healthy individuals using mazes of various difficulty levels and characteristics, I was able to gather, process, and examine gaze data to investigate patterns in eye behavior during decision-making problems.

Thesis Supervisor: Randall Davis
Title: Professor

Acknowledgments

Special thanks to my amazing advisor, Professor Randall Davis of the Human Computer Interactions Lab at CSAIL, and to Dr. Dana Penney, Neurology Specialist from Lahey Hospital and Medical Center, for guiding me through this project. I've spent countless hours looking up to you two for advice, opinions, materials, and so much more. Thanks for looking out for me, even in ways beyond your duties such as encouraging me to speak louder and more confidently. Your support has been invaluable and I have learned so much from working with you.

Thanks to LeeAnne Lipert, Neurology Clinical Research Associate at Lahey Hospital and Medical Center, who walked us through the training process for conducting the subject testing.

Thanks to Steve Diles, Stephen Li, and Andrew Wong for being wonderful and knowledgeable lab mates. Thanks for helping me with software problems, replying to panicked texts, and acting as sanity checks during stressful times.

Thanks to my parents and my friends for cheering me on and feeding me snacks and vegetables throughout the course of this project. Without your love and support, I know I couldn't possibly be where I am now.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Studying eye behavior can provide us with insights about the cognitive load an individual experiences when solving challenging problems. There are several types of behaviors that we can examine, discussed in detail below.

Throughout this project, I used various maze tests to conduct eye tracking studies on healthy individuals from the MIT community. I processed the data to remove noise, then analyzed and visualized the data to determine how eye behavior varied during decision making tasks.

1.2 Eye Behavior and Cognition

Alzheimer's disease is an irreversible neurological disorder that causes dementia, affecting many cognitive functions such as memory, reasoning, concentration, and decision making [13]. In severe cases, the patient can no longer function on their own and must depend on others for basic life activities. According to the National Institute of Health, symptoms of Alzheimer's usually begin to appear when individuals are in their 60s, and the disease is among the leading causes of death in the United States.

Research has shown that cognitive effort is related to the behavior of the eyes [4]. Blinking patterns, pupil size, and saccade speeds appear to be correlated with cognitive workload. For example, the blinking rate of an individual may decrease as their concentration increases for a more difficult task. Patients with Alzheimer's disease are known to show increased delay before reflexive eye movements and changes in the length of their fixations [16]. These differences, among other characteristics, potentially indicate a loss of executive functioning. Although it's still unclear whether the correlation implies causation, this relationship is still useful for searching for signs that mental function has changed.

1.3 Types of Eye Behaviors

1.3.1 Fixations and Saccades

Two basic eye behaviors are fixations and saccades [5]. Fixations occur when the eyes are locked onto a target for an extended period of time. The minimum duration necessary to qualify as a fixation is up for debate, but researchers often use a threshold between 100 - 200ms [3]. When analyzing gaze data, a fixation is identified as consecutive data points gathered in the same area within a small error. Examining the frequency and positions of fixation points can reveal where a subject is directing their attention and focus.

Saccades occur when the eye moves quickly from one fixation point to another. In order to classify as a saccade, the eye movement must exceed an angular velocity of $20°/s$ for at least 4ms [18]. In general, there is a "ramp-up" time, or latency, of around 200ms [10] before a saccade starts, so saccades that occur less than 20ms apart are combined as one due those latency expectations.

Saccade behavior can be correlated to the actions being performed. When a subject draws a line, their gaze path may consist of several saccades that jump from fixation point to fixation point along the length of that line. When solving puzzles, a sudden series of random or chaotic saccades might indicate that the subject is confused and scanning around the page for a solution, while a smoother sequence of saccades with fewer direction changes may indicate that the subject has learned how to perform a task.

1.3.2 Blinks

We are also interested in blink behavior. There are three types of blinks: spontaneous, reflex, and voluntary [1]. Spontaneous blinks occur automatically without any conscious effort or external stimuli, while reflex blinks, while not necessarily intentional, occur in response to stimuli such as bright lights or a sudden touch. Voluntary blinks are intentional.

Spontaneous blink frequency is correlated to mental performance, with lower blink rates corresponding to higher cognitive load [15]. Furthermore, the duration of spontaneous blinks has also been shown to be an indicator of increased mental workload [4]. One goal for the eye tracking study is to see whether the difficulty of certain drawing tasks has any effect on blink behavior.

1.3.3 Pupilometry

Pupil size and dilation can be surprisingly telling of cognitive processes: changes in pupil size on the order of < 1 mm can reveal changes in cognitive effort [7]. Therefore, we also examine the size of pupils to follow a subject's concentration.

Chapter 2

Equipment

2.1 Pupil

In order to track eye movements and gaze positions, we used the Pupil headset. This device is worn on the head in the same position that glasses would sit and is very lightweight (weighing about 1 ounce), minimizing intrusiveness. If a subject wears glasses, the headset can be fitted on top of them and secured by straps. The Pupil headset contains three adjustable cameras: two eye cameras are pointed to each of the eyes, using infrared sensors to locate the pupil, while a third camera (the "world" camera) is pointed outwards to the wearer's field of view. Together, these cameras create a model of the eyes based on the positions of the pupils in order to calculate exactly where the subject is looking in the world at any given moment [9]. The cameras take samples at a rate of 120Hz, but by polling the eye data from the two cameras, the software generates a sampling rate of 240Hz.

The Pupil headset cameras need to be calibrated for each use, even if the subject had tested twice in a row. This results because it is nearly impossible to position the headset in exactly the same way every time, and inaccuracies can result from the headset shifting around while the user is moving their head or bumping into other objects between trials. Gaze calibration involves asking the subject to focus on a round target (Figure 2-1). The target is moved around the table and the subject is asked to follow the target with only their eyes, while keeping their head and body

Figure 2-1: Circle target for gaze calibration.

as still as possible. The target is moved up, down, left, right, and across diagonals and the perimeter of the view of the world camera. More detailed instructions are available in the Appendix.

The software for the Pupil headset comes in two parts. The first, Pupil Capture, is used to record the videos from the three cameras on the headset. Various settings allow us to adjust the pupil detection sensitivity, indicate which test form is currently being used, and change video quality settings. The interface also displays timestamp markers that can be moved to indicate which portion of the video contains the calibration session we want to use, if multiple calibrations take place during a recording. Pupil Capture generates useful data such as pupil sizes, pupil positions, and whether it detects the fiducial marks of the paper forms we are using.

The second body of Pupil software is the Pupil Player. It takes the recording files from Pupil Capture and uses the pupil positions and surface detection data to calculate the positions of the gaze on that surface. The pupil data can be provided as either a 2D model, with the size of the pupils given in pixels, or as a 3D model, with the pupil size given in millimeters. The downside of the 3D model is that it is sensitive to noise and frequently produces unrealistic values. Hence, during the data analysis, 2D was the preferred model as the data was more accurate and faster to generate.

2.2 Pen

We had test subjects solve mazes using the Anoto pen, which automatically records all pen movements as timestamped coordinates that can later be uploaded to a computer. The test forms are printed with a faint and closely spaced set of dots which are viewed by a camera in the barrel of the pen to determine its position on the page. This is done at a rate of 75 samples per second and can determine the pen position to within 1000th of an inch.

The benefit of using the Anoto pen over a regular pen is that it provides a dynamic picture of the pen movement. Timestamps are given for every point, allowing us to know how fast the pen is moving at all times. Another benefit is that because the pen results are uploaded and backed up, we don't have to worry about losing the paper test after the subject testing is done.

The software used for the Anoto pen is called THink. THink allows us to view and score the pen drawings on tests such as the maze tests, symbol-digit tests, and clock drawing tests. For my research, I focused only on the maze tests.

2.3 Mazes

The test mazes we have currently come in three difficulty levels: easy, intermediate, and advanced. The difficulty of the mazes depends on the complexity of choices, path length, and frequency of embedded decisions (i.e. when a choice leads to more choices that have to be evaluated before the subject can be confident whether the initial choice was wrong). Each maze is designed to be balanced, meaning that the path that the solution takes is spread out to spend a relatively equal amount of time on both the left and right halves of the maze, as well as a relatively equal amount of time on the top and bottom halves of the maze.

On each paper form, there are two large mazes and one much smaller maze (see Figure 2-2). This very simple maze consists of only two horizontal lines, solved from left to right, and is meant to be used for speed calibration at the start of every maze

Figure 2-2: Small speed calibration maze, done before either of the larger mazes.

test.

For the larger mazes, there are two types (Figure 5-4). The first type, which we call the “choice maze”, resembles a traditional square maze with one correct solution path from the start to the finish and many junctions where the wrong decision leads to dead ends. The “choice” refers to the presence of these decision junctions: the subject has to make a choice as to which direction they should go. The other type of maze is the “no choice maze”. The no choice maze is nearly identical to the choice maze, except that extra walls are drawn at each decision junction to block off incorrect decisions. The result is that every position has exactly one direction forward for the subject to take; there is no way for them to take the wrong path if they do not backtrack or cross through walls. In both the choice and no choice mazes, the solution paths begin at the bottom of the maze and end at the very top.

Each maze is about 5 inches × 5 inches in size, and each is printed on half of an 8.5 × 11 inch paper form. The small calibration maze is printed below the no choice maze. For every maze test, the subject is given the paper form folded in half such that only the no choice and the calibration mazes are visible. At the start of each maze test, the subject is first asked to draw a line through the calibration maze. Then, the subject is asked to solve the no choice maze. Once completed, the form is flipped over to reveal the choice maze, which they are then asked to solve. Only one large maze is visible at a time. The purpose of having these two types of mazes is to form a baseline for behaviors such as how fast the subject draws their solution in the presence and absence of choices and how their eyes move while doing so. We want to see how these behaviors change when cognitive load is introduced.

In addition to the mazes, test subjects were also asked to solve a Symbol Digit (SD) test after completing the mazes. Another researcher in our group was responsible for analyzing these SD tests, which involve puzzles asking the subject to match numbers to either symbols or other numbers.

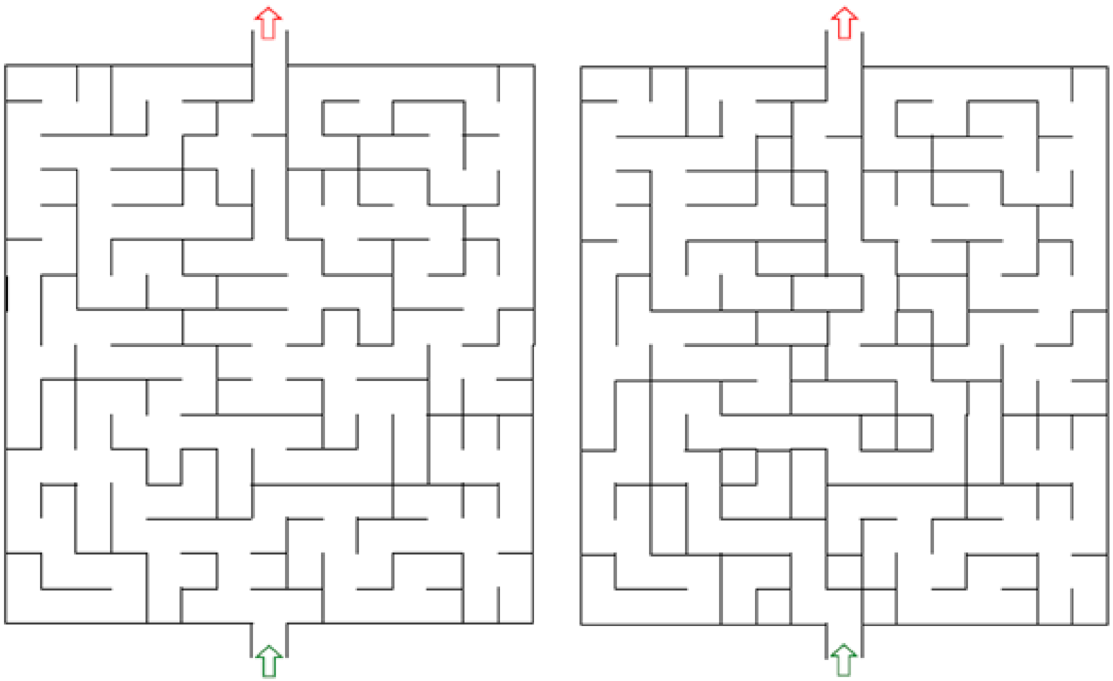


Figure 2-3: Sample choice (left) and no-choice (right) mazes.

Chapter 3

Subjects

3.1 Training

Prior to conducting tests on any people outside of the lab, we submitted our project to the Institutional Review Board (IRB) and received approval to proceed with the study. Following that, every test giver was given training for how to properly conduct the tests. For example, we need to ask subjects to keep their pen on the paper during the maze test and avoid lifting the pen when thinking or after a mistake. Instead, the subjects are asked to backtrack to where they last made the mistake. We also learned about good practices to follow when testing subjects, such as how to make them feel at ease, remove potential distractions, and protecting the privacy of the subjects.

3.2 Recruitment

The eye tracking tests were conducted on 60 healthy members of the MIT community, between the ages of 18-30s. The subjects were recruited after they responded to an email advertisement or after hearing about the study from a prior participant.

Initial tests were conducted on volunteers to give us a sense of how long the overall testing process would take. Once we were familiar with the process, an email advertisement was sent out to the CSAIL mailing list. Subjects were offered compensation of a $20 Amazon gift card. Twenty volunteers responded, primarily grad students.

The following week, the advertisement was forwarded to an undergraduate psychology class, with a fantastic response rate. From those respondents, the first thirty undergraduates were recruited to participate in the study. It may be interesting to note that the largest demographic representation among the test subjects were Asian female students, who made up half of the total number of subjects.

3.3 Testing

Testing took place in multiple conference rooms throughout the CSAIL building over the course of three weeks, varying due to scheduling availability. We avoided rooms with windows, and if a room did have them, the blinds were pulled down to remove outside distraction and natural light changes. Before every test subject, we measured the light intensity in the room using a light meter.

Each study took roughly 30 minutes to conduct. To reduce stress for both the subjects and the testers, each subject was given a 45 minute timeslot. Prior to testing, subjects were given a consent form to read over and sign.

The subjects were then given five tests in order. The first three tests were all maze tests, with three different levels of difficulty (easy, intermediate, and advanced) given in a randomized order. For each maze test, the subject was asked to first solve the calibration maze, then the no-choice maze, and finally the choice maze. Following the maze tests were two identical SD tests. Throughout the test, the subjects were not told which maze level they were taking, nor were they informed that the two SD tests were the same. The testing procedure is described in more depth in the Appendix.

Chapter 4

Data Processing

As mentioned earlier, eye tracking data is known for being extremely noisy. Little jitters in the signal could be caused from anything from vibrations when the user moves their head to inaccuracies in measurement or noise from the equipment. These jitters could potentially be confused with meaningful eye movements such as saccades. To separate the noise from the meaningful data, we need to understand what behaviors are physically possible for the human eye.

4.1 Properties of Eye Movements

The first item to look at are saccade properties. According to Findlay [6], the typical duration of an eye saccade is between 20ms - 100ms. The duration varies according to the angular change of the eye. The raw gaze signal often contains small periodic patterns that repeat at a high frequency (Figure 4-1). If this frequency is higher than 50Hz (i.e. more frequent than once every 20ms), then this measurement is most likely to be noise. During data cleaning, I used 60Hz as my cutoff value for classifying noise.

Additionally, we know that the angular velocity of the human eye has a speed limit of $900°/s$. We can translate this to how fast we expect the gaze point to travel on the paper surfaces. We assume that a subject holds their head about 15 inches above the form. The camera sampling frequency is 120Hz, but due to offset sampling times of the two eye cameras, the gaze may be sampled at what appears to be 240Hz.

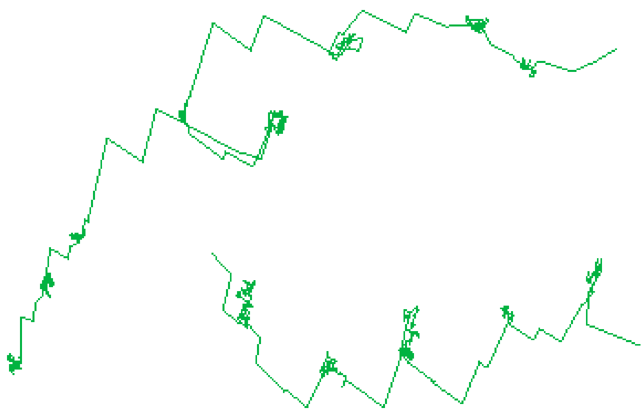


Figure 4-1: Two segments of unfiltered data. The periodic peaks are too fast to be physically possible for the eye and are most likely just noise.

Therefore, the time between samples is 4.16ms:

$900°/\mathrm{s} \cdot 0.00416s = 3.75°$

$7.5° \cdot (\pi/180°) = 0.065$ rad

0.065 rad $\cdot$ 15 inches = 0.98 inches

Therefore, a distance differential greater than one inch in 4.16ms would be physically impossible for the human eye, and any occurrences would require us to take a closer look at the data and recordings to figure out what had happened.

Furthermore, saccades do not occur instantly. As mentioned in Chapter 1, there a latency of around 200ms [10] before a saccade starts. This translates to around 24 gaze samples. Therefore, any “saccades” with a frequency higher than 5Hz are also physically impossible.

4.2 Data Cleaning

Now that we know how to identify noise, we need to clean it from our data. This can be done with a low-pass filter.

Designing a low-pass filter for this gaze data requires several considerations. First,

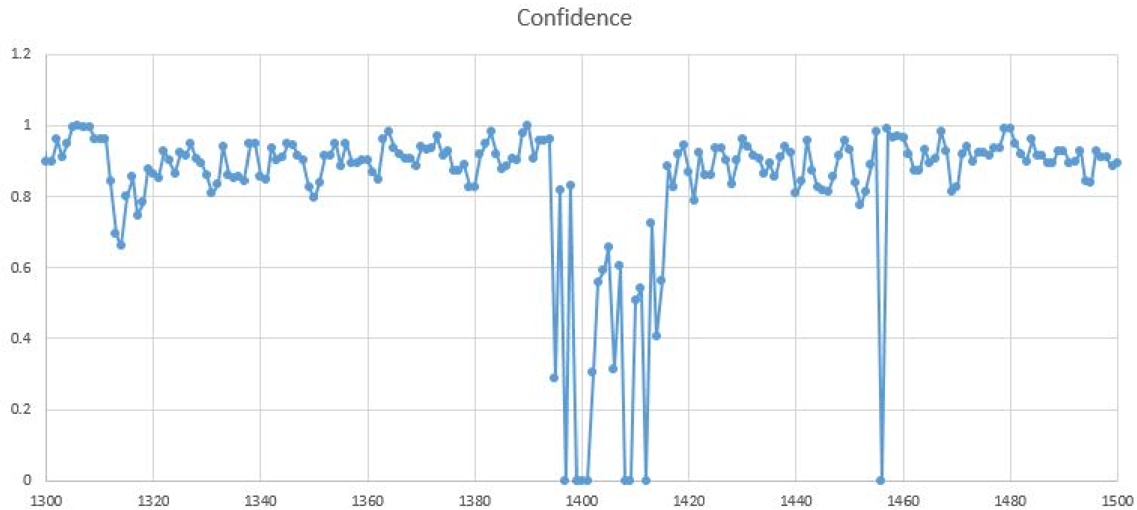


Figure 4-2: Detection confidences values around a blink. The confidence values decrease to 0% before rising back up when the eye reopens.

we know the range of speeds and frequency of saccades, so we can safely filter out any portions of the gaze data that are too high frequency. We will use a Butterworth filter to ensure that we have the best preservation of frequencies below the threshold[1].

Second, there are some data points that don't correlate to motion - when a subject blinks, the surface position coordinates can't be properly calculated and as a result are set to values that are unrealistically large in magnitude (such as being hundreds of inches away). Fortunately, there are indicators for whether a blink has occurred: there is a sequence of gaze samples for which the detection confidence values gradually decrease to (or close to) 0% and then gradually rise back up (Figure 4-2); during this time, the calculated positions will no longer be within the dimensions of the surface (i.e., will no longer be on the form), which Pupil may indicate with an "on-surface" parameter set to False for each sample. After the artifactual blink positions have been removed, we smooth out the signal through the use of a median filter, which also behaves as another low pass filter.

Finally, there are times when the subject is not blinking but the gaze position still

[1] Many types of low-pass filters may have response "ripples" in their passband, which can potentially distort or add more noise to the signal being filtered. Compared to other low-pass filters, a Butterworth filter has a very flat response, meaning that we are less likely to distort the portions of the signal that we wish to keep.

leaves the surface, such as when the subject looks away to glance at something in the room during the test. If the subject glances just a little bit to side of the form, the maze surface is still detected and Pupil will provide a position with values that are outside of the dimensions of the surface but that are still realistic in terms of how far away they are from the page. However, if the subject looks around the room, the surface may no longer be detected and Pupil will generate position values that are large and unrealistic, just like with blinks. The difference is that when looking away, the confidence of pupil detection does not necessarily decrease, and the time that the gaze is off the surface is much longer (such as 5 seconds). Oftentimes it helps to look back at the video recordings to confirm what activity had actually occurred. Hence, distinguishing between blinks and glances in our gaze data requires mathematical checks with help from manual review.

4.3 Visualization

In this paper, the maze backgrounds will be replaced by a blank maze (Figure 4-3), as information about the maze forms that we used for subject testing is not yet available to the public. Most of the visualization is done using Plotly, which allows us to create interactive graphs and plots. As we can see, the gaze signal largely follows a discernible path through the maze, although the signal is too chaotic and spread out for us to see the details of specific paths and turns. This also is due to the fact that during the tests, subjects frequently look ahead or around the maze beyond just the solution path that they are following, in both the choice and no choice mazes.

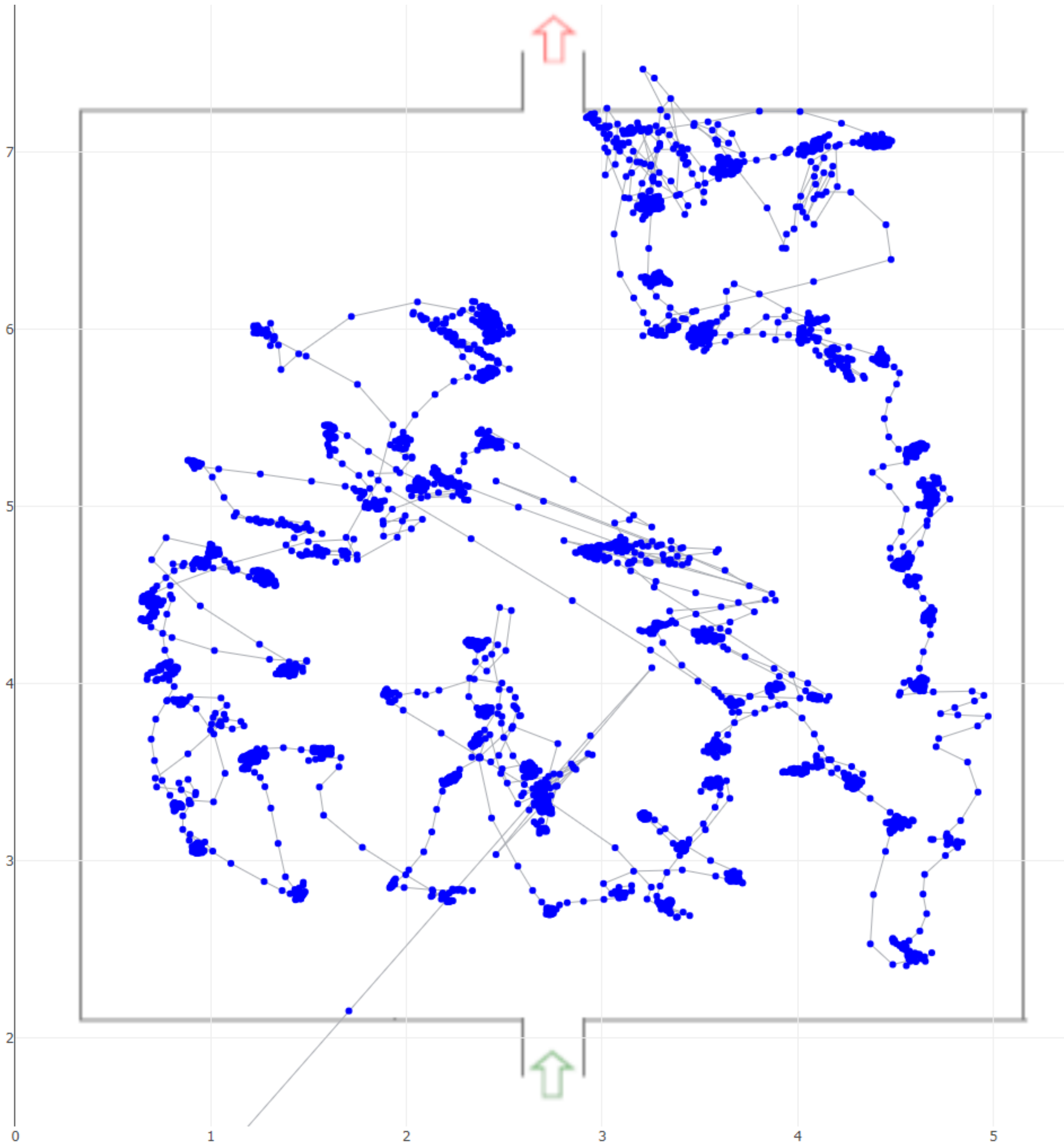


Figure 4-3: Sample gaze data path in a blank maze, with walls removed.

Chapter 5

Analysis

5.1 Blink Patterns

One of the first things I looked into was the blink behavior across different mazes. To detect blinks, the Pupil Player performs a convolution of the gaze data with a gradient filter, consisting of a $1 \times n$ array where the values in the first half are -1 and the values in the second half are 1:

$$[-1, -1, ... -1, 0, 1, ..., 1, 1]$$

This gives us the gradient of the eye detection confidence over time, as calculated within a window of length n. The value of n is however many samples can be found within a 200ms time window. Given our sampling frequency of 240Hz, this gives us 48 samples, assuming there is no missing data. To determine whether a blink occurred, we check whether the change in smoothed confidence fell below a fixed confidence threshold (set to 50% in the default case) for a few milliseconds before rising back up.

However, the algorithm used by Pupil did not take in account the realistic timespan of a typical blink. According to several papers, the expected length of a blink, from the moment the eyelid begins to drop to when it is back up, lies within 100-300ms [17] [2]. With the Pupil algorithm, even a lengthy drop in confidence (e.g. 1

Subject	Difficulty	NC Blinks	C Blinks	% Change
1	Easy	8	6	-25
1	Advanced	4	6	50
2	Easy	1	1	0
2	Advanced	5	3	-40
3	Easy	37	21	-43.2
3	Advanced	21	13	-38.1
4	Easy	6	6	0
4	Advanced	5	8	60
5	Easy	0	0	-
5	Advanced	3	1	-66.7

Table 5.1: Table of some blink counts across five different subjects. “NC Blinks” is the blink count over the No Choice maze, while “C Blinks” is the blink count over the Choice maze.

second), such as may occur when the subject looks off the page or when pupil detection is temporarily lost due to extreme angle, will still be counted as a blink. The lack of duration considerations also results in the algorithm detecting even a brief drop in confidence, due to measurement noise, and also counting that as a blink. In some extreme cases, Pupil would tell us that a particular subject had blinked as much as 300 times in a one minute interval, which is physiologically improbable.

To deal with this issue, I wrote a script that took timing into account and used that instead of Pupil's blink detection algorithm. Similar to the Pupil algorithm, I used the drops and rises in confidence values to isolate potential blink periods. I then calculated the range on the number of samples that realistically span the length of a normal blink (24 to 36 samples), and filtered out all those potential blinks with durations that lay outside of those ranges.

According to Coral [4], the frequency and duration of eye blinks can indicate how much concentration an individual is exerting on a problem. More specifically, the higher the level of concentration and focus, the fewer the number of blinks we expect.

Some sample blink data is presented in Tables 5.1 (blink counts) and 5.2 (blink rates). Blink rates are calculated by taking the total number of blinks over the time it took the subject to complete the maze test.

We can also visualize the blinks on the gaze data. While the blink positions are

Subject	Difficulty	NC Rate	C Rate	% Change
1	Easy	0.15	0.14	-10.6
1	Advanced	0.12	0.16	43.7
2	Easy	0.02	0.02	20.8
2	Advanced	0.06	0.05	-24.1
3	Easy	0.71	0.47	-32.8
3	Advanced	0.51	0.37	-27.1
4	Easy	0.03	0.03	-0.9
4	Advanced	0.14	0.23	64.8
5	Easy	0	0	-
5	Advanced	0.06	0.01	-82

Table 5.2: Table of some blink rates (in blinks per second) across the same five subjects as in Table 5.1. "NC Rate" is the blink rate over the No Choice maze, while "C Rate" is the blink rate over the Choice maze.

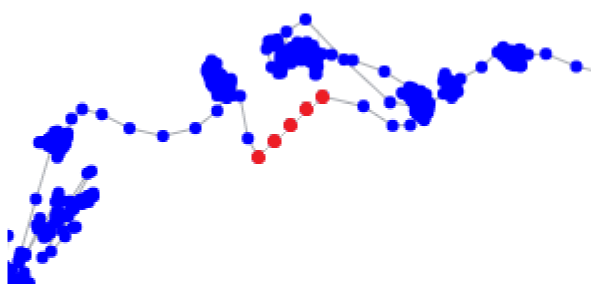


Figure 5-1: Example visualization of a blink (in red) in part of the gaze signal.

invalid, the gaze positions before and after those data points are still on the maze surface, so we can extrapolate a line between those neighbors and color them to stand out. The number of points on the linear extrapolation segment should correspond to the number of samples for which the blink takes place.

5.2 NSLR Segmentation

Since most saccade movements can be approximated as linear paths between the static fixation points, a gaze path could be represented as a set of line segments, connected with fixations as vertices. One method of separating a gaze path into linear saccades and fixations is Naive Segmented Linear Regression (NSLR), a segmentation algorithm that relies on a Hidden Markov model classifier to predict which movement type a gaze data point is most likely to be.

NSLR is capable of classifying four different types of eye movements: saccades, fixations, post-saccadic oscillations, and smooth pursuits. Post-saccadic oscillations refer to the small corrections the eyes make after transitioning from a saccade to a fixation, while smooth pursuits describe the movement of the eye as it follows a moving target. However, because our drawing tasks do not involve any moving targets other than the pen, we will mainly focus on the first two categories.

This segmentation also serves the additional function of denoising the gaze signal, as an alternative to the signal processing low pass filtering. However, the downside is that NSLR assumes the saccades in the gaze data are all linear, which may not always be the case. Furthermore, the algorithm was sensitive to small changes in the data. By cropping out the last 100 data points and rerunning the algorithm, the newly generated fixations points would have all shifted to slightly different locations.

Details of the NSLR algorithm are documented in [14]. I adapted a version of the code provided in their paper and ran it on some of the maze tests, with a sample result shown in Figure 5-2, which shows the simplified line-segment representation of the saccades and fixations produced by the algorithm.

5.3 Fixations and Gaze Speed

An alternative approach to identifying the fixations in a gaze signal is to simply look into the speed of the gaze throughout different parts of the maze test. During fixations, the eye is focused on a single position and is therefore has speeds that are close to 0. During saccades, the eye accelerates to much higher speeds, reaching a maximum angular acceleration of 900°/s, as discussed in the previous chapter.

Using our gaze data, speeds were calculated for each point using three gaze positions, that of the current gaze sample as well as that of the two timewise adjacent points before and after that sample. The sample points were then colored based on their speeds using a color scale. In the visualization in Figure 5-3, the fixations are

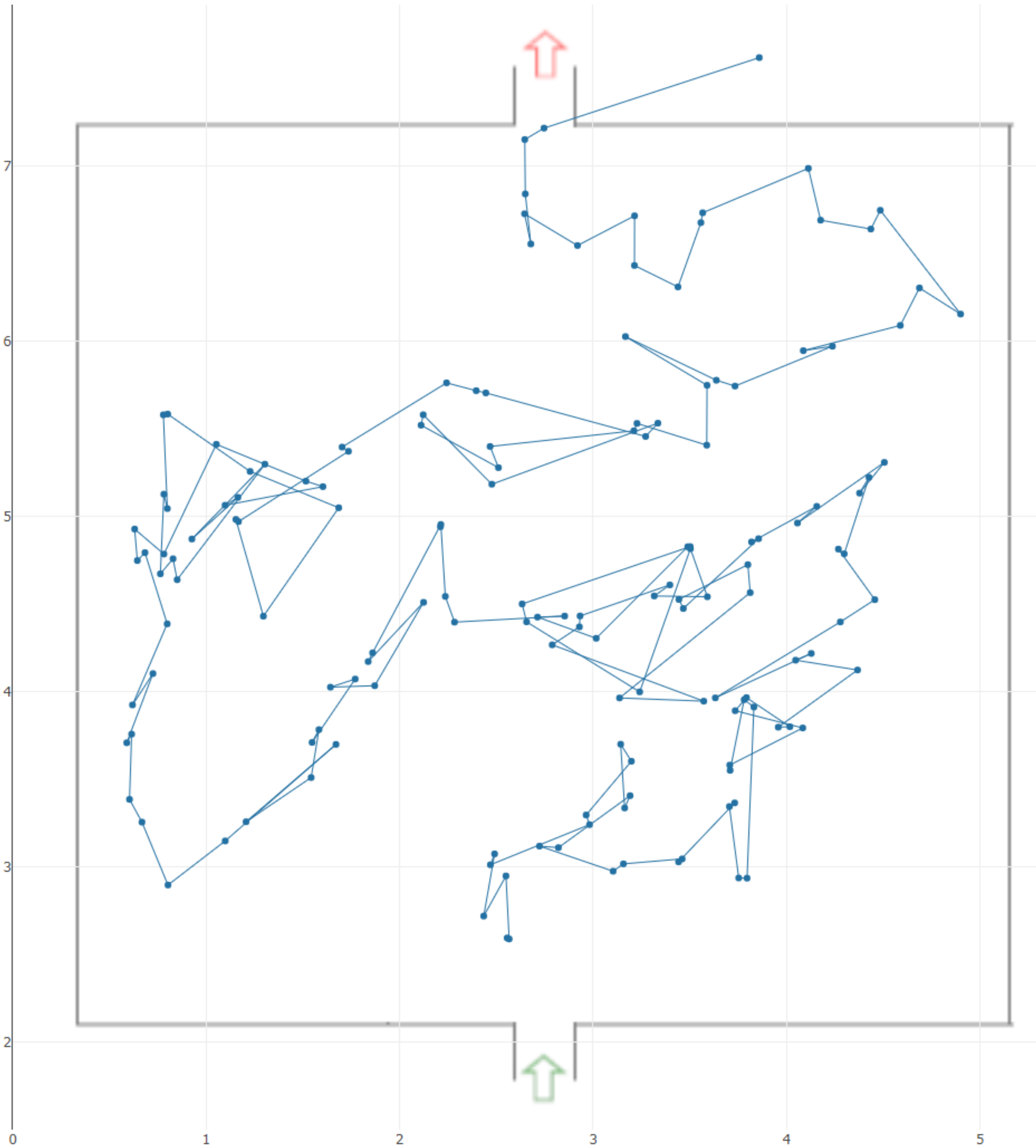


Figure 5-2: NSLR segmentation of gaze data.

depicted in dark purple, while any gaze points with speeds higher than 3 inches/s are colored in yellow. As we can see, both NSLR and the gaze speed method were able to capture many of the same fixations points.

5.4 Decision Areas

A major topic of interest is how the eye behaves when encountering decision areas in the mazes, i.e. junctions in the choice mazes where the subject has to decide which direction to continue in (in the corresponding no choice mazes, the extra options are sealed off with a border.)

One might assume that the decision areas in the choice maze will always require more cognitive effort due to the presence of choices, but it's possible that the shapes formed by the sealed off areas in the no choice maze may add cognitive load by being distracting, especially if they are small and rectangular (Figure 5-4), which one would not expect in a typical maze puzzle. Subjects may also be confused if it is their first time solving a no choice maze.

The following two sections discuss two ways I examined eye behavior around decision areas - through pupillometry and the distance between gaze and pen tip.

5.5 Pupil Size

Variations in an individual's pupil size throughout a challenging task can be indicative of changes in cognitive focus (such as higher levels of concentration or confusion). Hence, I investigated how pupil size varied throughout different maze tests.

The range of normal pupil sizes among healthy individuals is between 2-8mm [12]. However, this range covers extreme cases, with 2mm occurring in very bright blinding light while 8mm occurs in nearly pitch black darkness. During our subject testing, we measured the light intensity in the room before each test. Hence, in a typical room lighting of 150-550 lux, the expected pupil size is closer to 3.5-5mm [11].

Pupil Player can generate 2D pupil radiuses, given in pixels, for every gaze point.

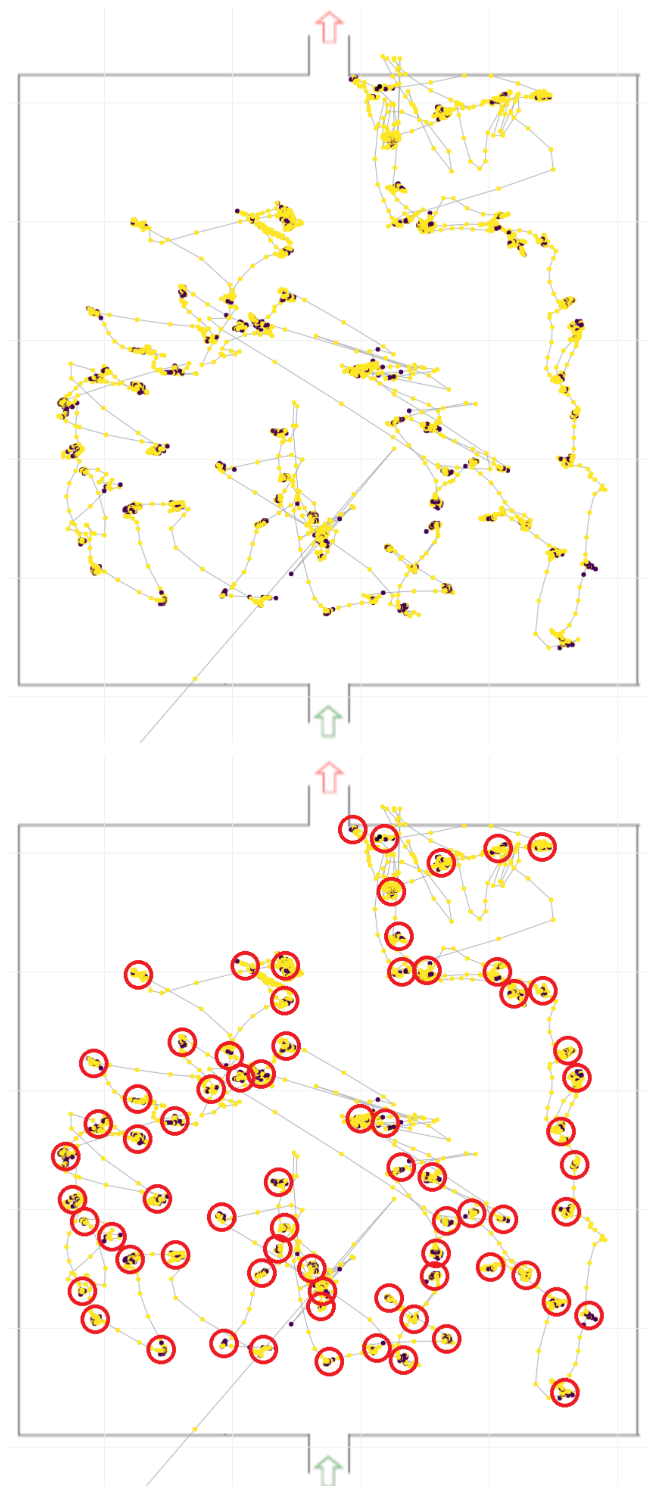


Figure 5-3: Fixations (purple) as identified by gaze speed on the surface. Fixations circled in second image.

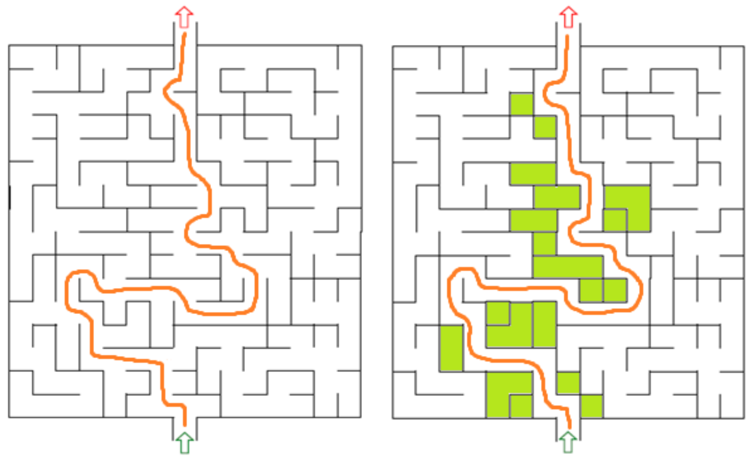


Figure 5-4: Sample maze with solution pen paths drawn in. In the no choice maze on the right, the sealed off junctions sometimes form small unexpected shapes that could potentially be distracting to a subject.

The range of the pixel values varied from person to person, depending on the distance between the camera and the eye as everyone has a different facial structure. As a sanity check, we can calculate what we might expect to be a reasonable pixel size range.

When I wore the Pupil headset, cameras sat 1.5 inches (38.1mm) from my eyes. We can assume that the monocular horizontal field of view is about 90° [8], which is also what is listed in the Pupil settings. The resolution of our camera is set to 1280×720. Looking at each half the field of view (at 45°), we calculate that at the distance of the eye, one pixel covers $38.1 \cdot 2\tan(45°)/1280 = 0.060$mm. To translate this into pixel ranges:

$$3.5\text{mm}/0.060 = 58.8\text{mm}$$

$$5\text{mm}/0.060 = 84.0\text{mm}$$

The subject used in most of our examples has a pupil pixel range of 65-85mm. When doing multiple tests on myself, my pupil sizes lie in the 70-100mm range. Given the math above, we see that these values are realistic for the human eye.

When the gaze points are colored based on the pupil size, clear clusters can be seen in many of the maze tests (Figures 5-5 and 5-6). Larger pupil sizes tend to appear near the beginning or end of mazes, or around areas with many decision points.

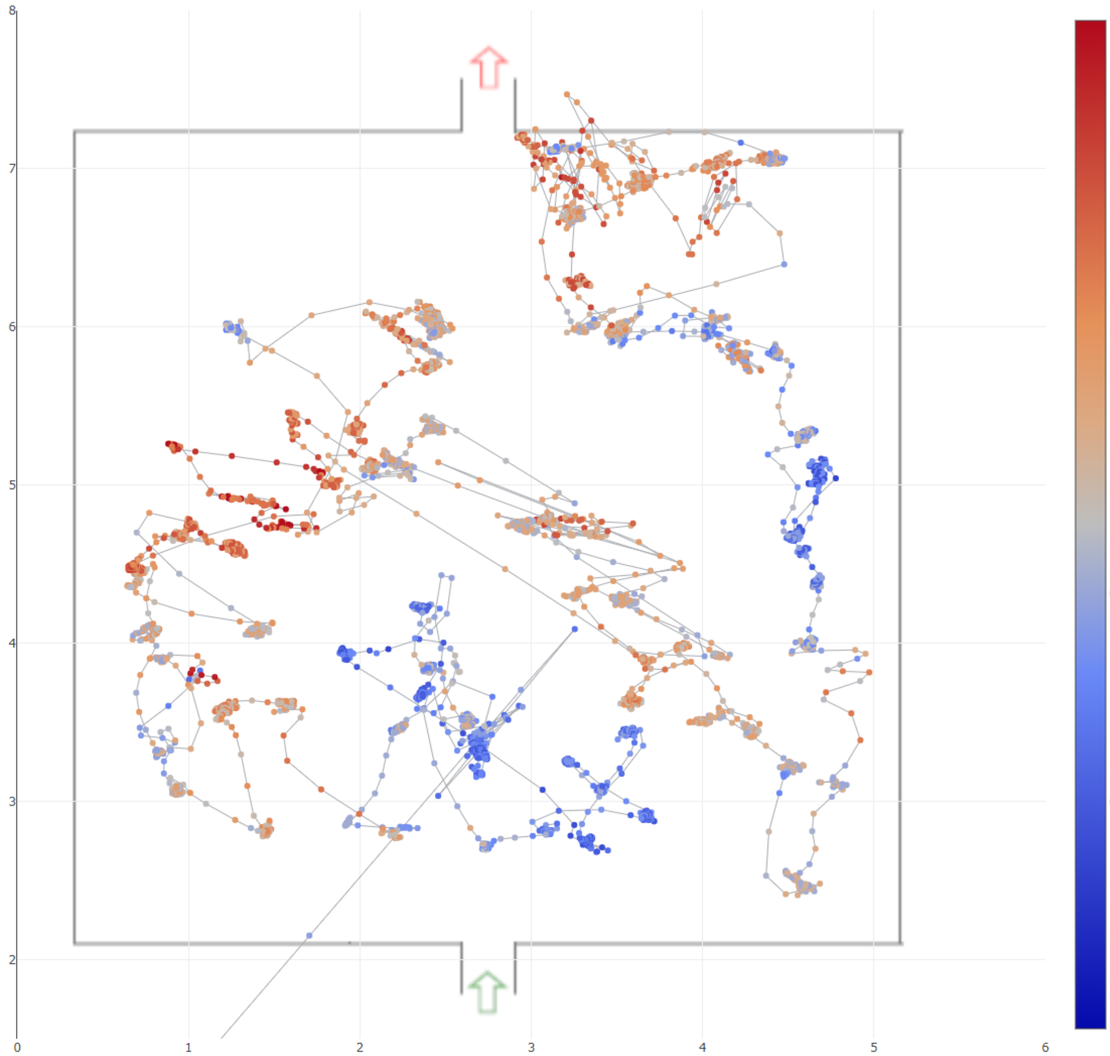


Figure 5-5: Gaze points colored by pupil radius (in pixels). Larger radius sizes are in red, while lower radius sizes are in blue. Clusters of colors can be seen, such as the blue area at the beginning of the maze and the red areas at the end and on the left.

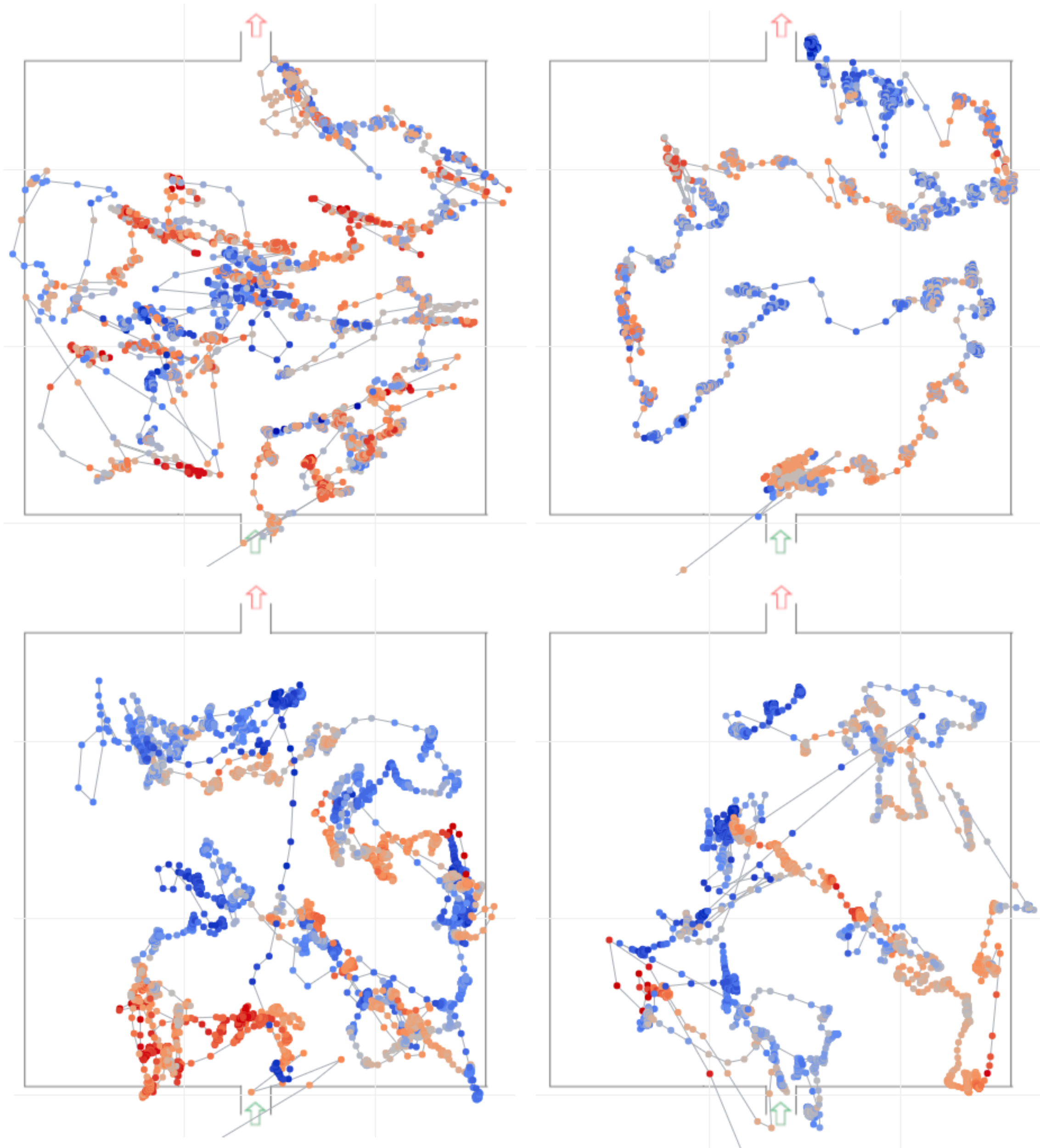


Figure 5-6: More maze samples where the gaze points are colored by pupil radius. Larger radius sizes are in red, while lower radius sizes are in blue. Each row of mazes are conducted by the same subject, with the choice maze on the left and the no choice maze on the right.

More excitingly, in the majority of the mazes, the median of the pupil sizes were higher in the choice mazes than in the no choice mazes. As higher pupil radius implies more cognitive load [4], this result is aligned with our expectations since the choice mazes are designed to be more challenging. In the Figure 5-7, individual squares in the mazes are colored according to the value of the mean of the pupil radiuses of the gaze points contained within those squares [1]. The means are compared around the median of all radius sizes such that radius sizes greater than the median are colored red and radius sizes less than the median are colored blue. This was done to investigate whether areas of higher or lower pupil radius sizes correlated to any decision areas.

People may work differently, so we used each individual as their own control. We take the difference between choice and no choice mazes of the same difficulty from the same individual by subtracting the means of corresponding squares. As seen in Figure 5-8, the differences, when each point is present, showed that this particular individual's pupils were larger on average throughout the choice maze, as indicated by the number of red squares and the intensity of the colors.

5.6 Gaze/pen distance

For the majority of the tests, the pen generally very closely follows the correct solution path in all the mazes. The gaze, on the other hand, may deviate from the solution path when the subject is scanning around if surprised, confused, or thinking ahead. From the gaze data of several tests, the eye movement tends to be more erratic around decision squares, as the subjects must look around to decide which route is the best one to take.

I initially tried plotting the difference in position between the gaze and pen points over time, as seen in Figure 5-9. The red bars indicate times when the current pen data was within a decision square. There are several peaks which occur in or near a

[1] At 240 samples/second, there are numerous measurements made while the gaze is within a square.

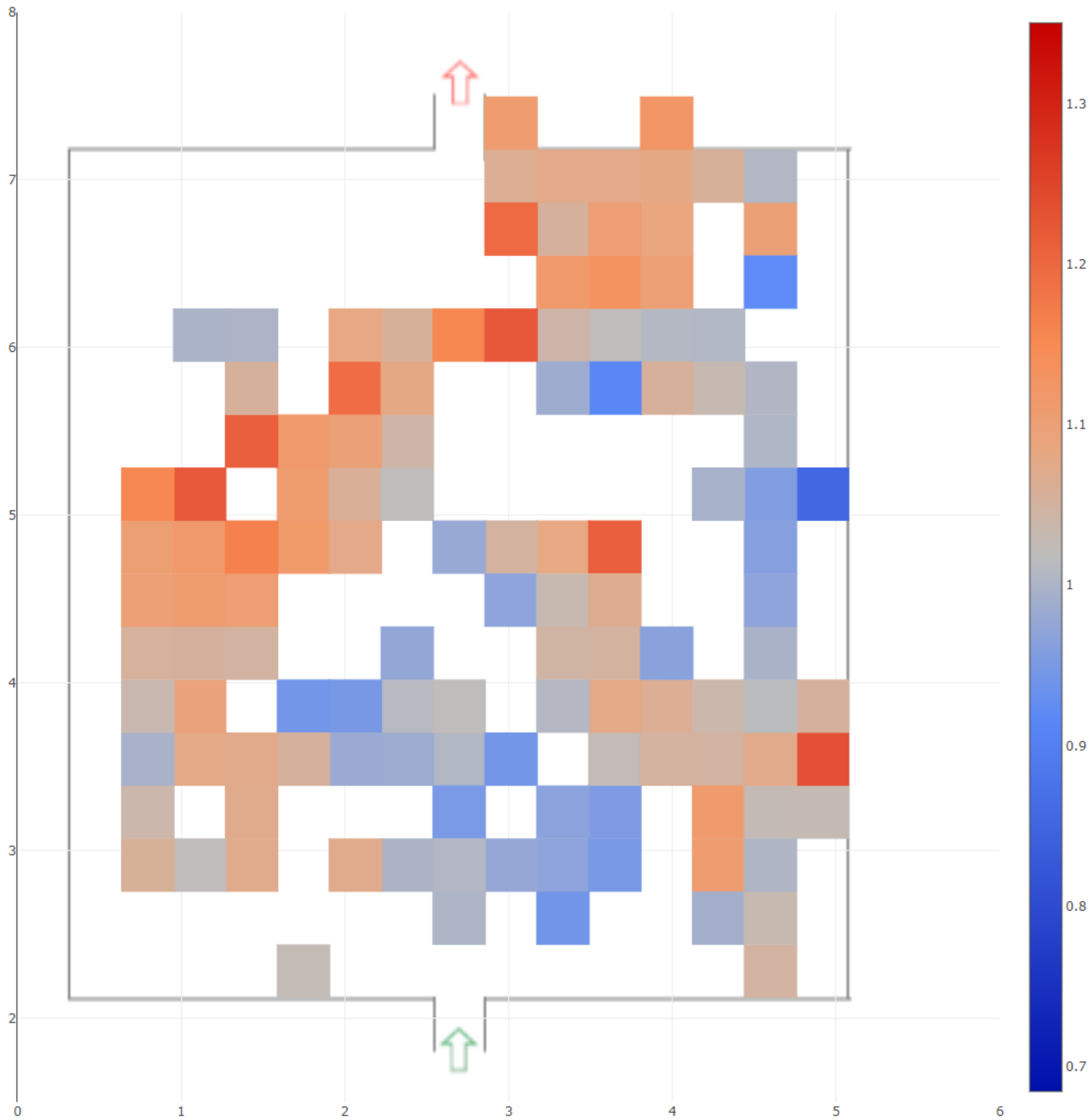
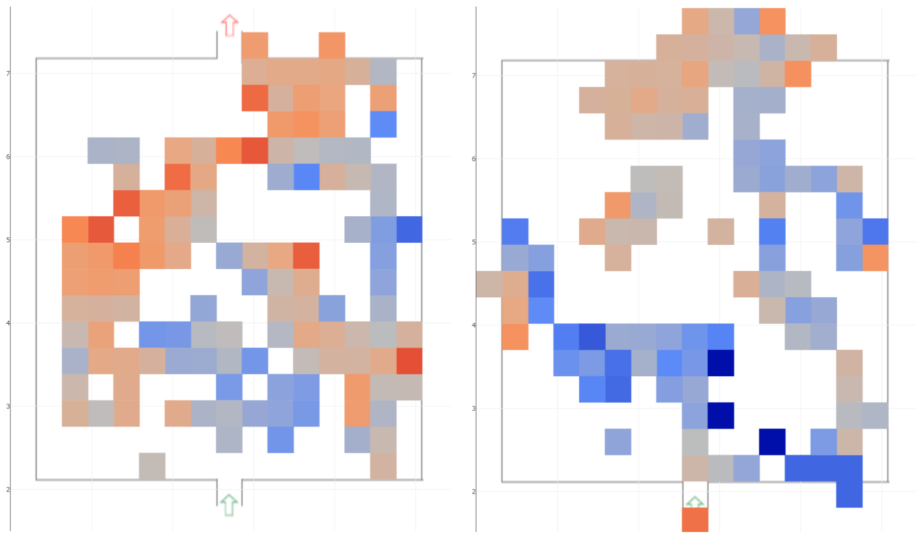


Figure 5-7: Heatmap of pupil radius means within individual squares on the choice maze. The scaling is normalized about the median of all radius values.



(a) Choice maze (b) No choice maze.

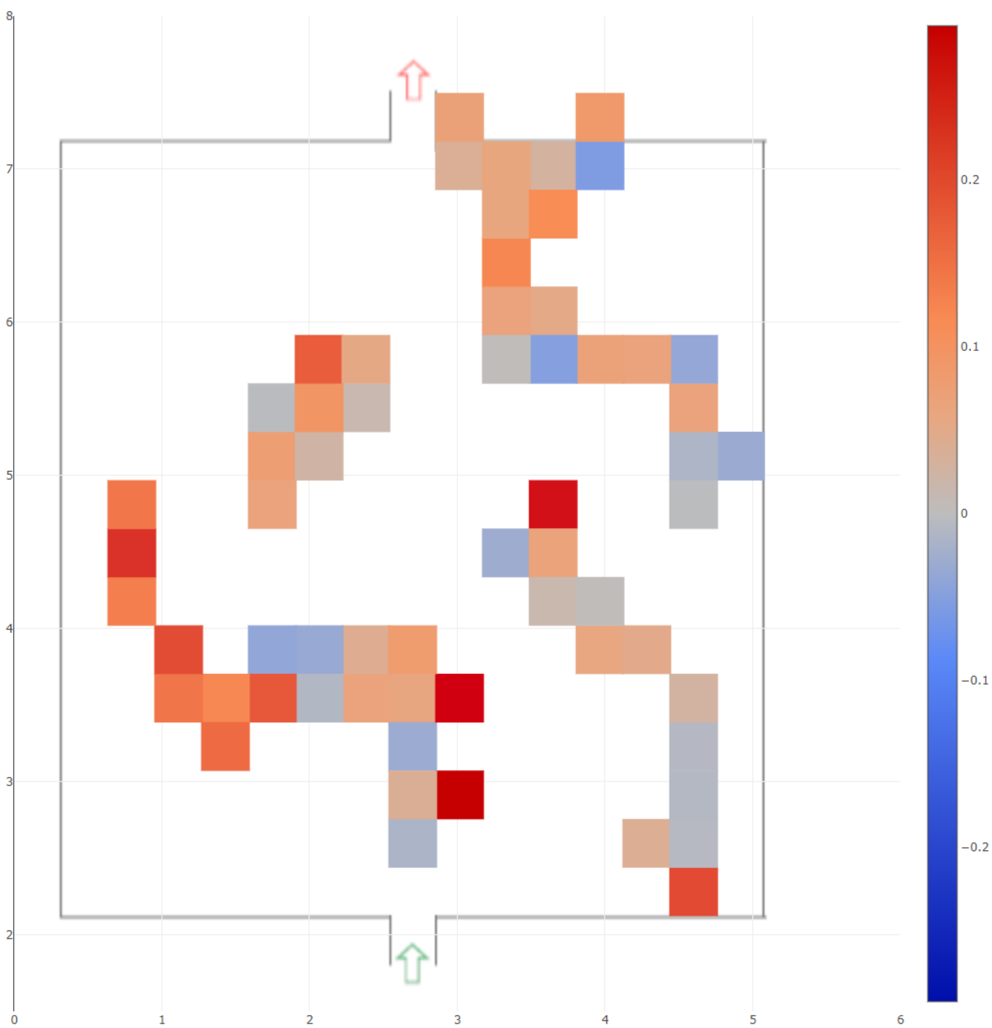
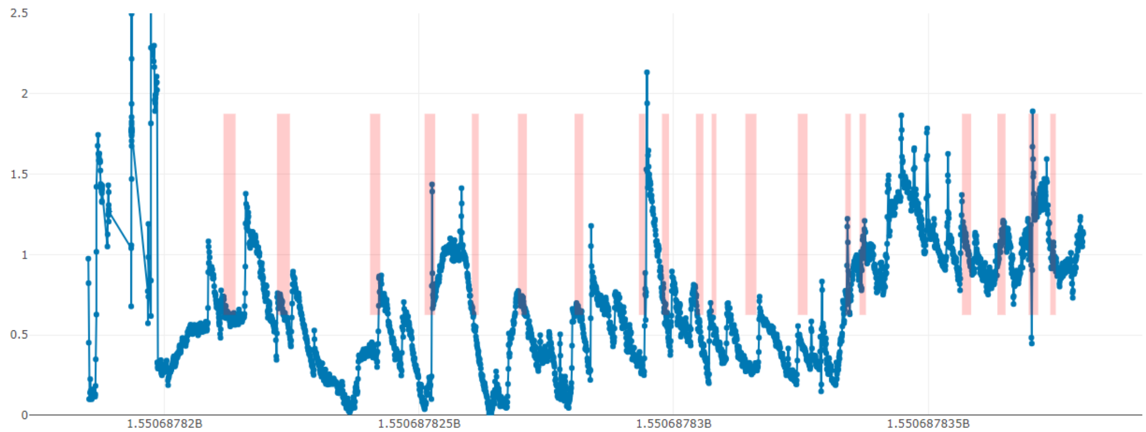
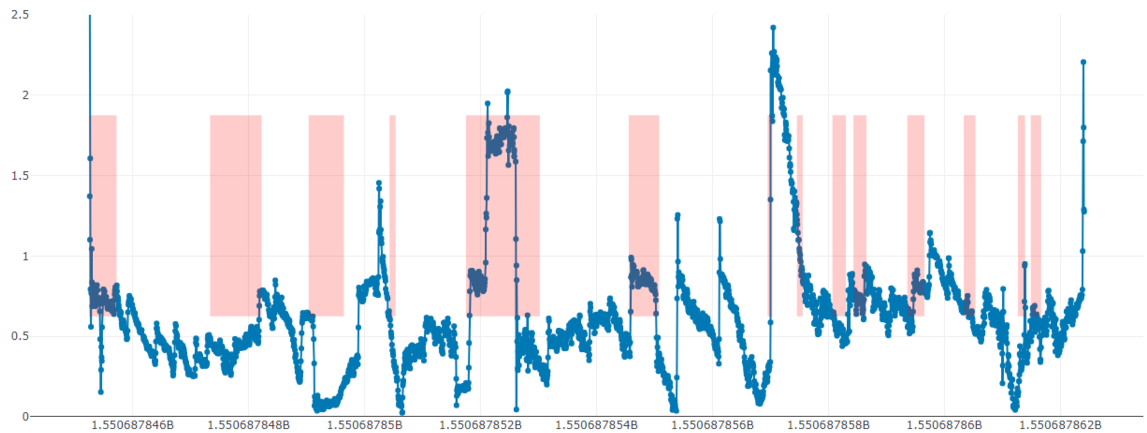


Figure 5-8: Difference of the heatmaps of the choice and no choice mazes.



(a) No choice maze.



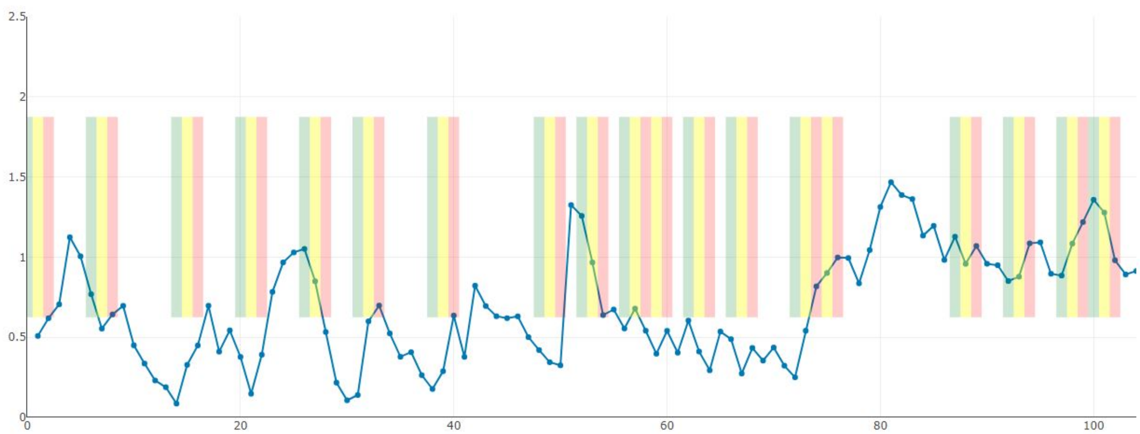
(b) Choice maze.

Figure 5-9: Difference in distance between pen and gaze over time.

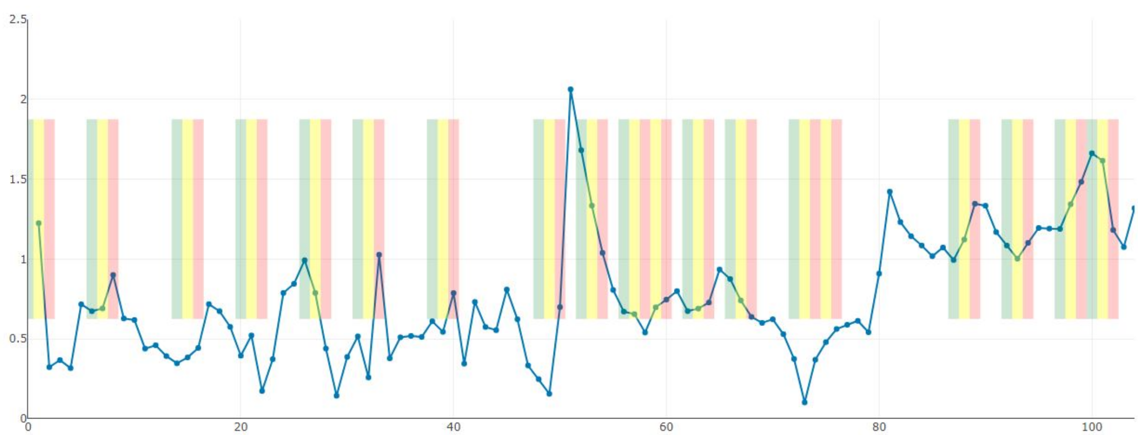
decision region in both the choice and no choice mazes.

However, the problem with the previous representation is that because subjects may potentially take vastly different amounts of time on different parts of corresponding choice and no choice mazes, its difficult to compare the two based on location along the solution path. This can be resolved by taking the mean of the distance difference within squares along the pen path, similar to as was done previously for the pupil radius. In Figure 5-10, the two mazes are aligned by location. The indices on the x-axis show how many squares along the solution path the pen is current at. Colored bars are also drawn to show whether the current position is in (red) or is near (yellow and green) a decision square. As we can see, the mean distance difference had a higher maximum value in the choice maze.

It's interesting to note that the peaks tend to occur in similar locations along the solution path - both have high value peaks at index 51, 81, and 100, indicating that the subject scanned around during those positions regardless of whether there were decisions or not. Looking back at the maze, those three indices correspond to squares that occur right before a long vertical or horizontal segment ("hallways"), perhaps revealing that the subject looked up those hallways to determine the total length or to see what the maze structure was at the end.



(a) No choice maze.



(b) Choice maze.

Figure 5-10: Mean difference in distance between pen and gaze over distance by the squares of the solution path. Red indicates a decision square.

Chapter 6

Discussion

6.1 Results

As expected, we found evidence of subjects exerting more cognitive effort for the choice mazes as compared to no choice mazes of the same difficulty level.

From looking at the pupil sizes, we see that subjects exhibited changes in pupil size throughout the mazes that often occur suddenly and in relation to a specific geographical location on the maze. Although the patterns varied, the majority of subjects showed a dramatic change in pupil size after the start of the maze as well as right before the end. Clusters of colors can also be seen throughout the gaze paths. More work will need to be done to determine the relationship between these clusters and the features of the mazes.

Furthermore, when comparing the size of the pupils between the choice and no choice mazes, we found that the median of the pupil radius was consistently higher in the choice mazes. Since pupil size generally increases as attention increases, this confirms our belief that subjects have to concentrate more for the choice mazes. Additionally, when comparing mean pupil sizes in corresponding squares along the paths in the choice and no choice mazes, we found that the mean pupil size in each square was also more likely to be higher in the choice mazes.

From examining the distance between the pen and the gaze at corresponding timestamps, we can see how much a subject's gaze moves around throughout a test.

Often at a decision junction in the choice maze, a subject will look around and exhibit more chaotic gaze behavior as they try to determine which decision is the best. Peaks of maximum distances occurred during or before decision squares in the choice mazes, but also before unusual features in the mazes, such as long vertical or horizontal hallways. In particular, when a peak occurred before a hallway in the choice maze, there would also be a corresponding peak in the no choice maze in the same position. This may show that visual features of the mazes can prove just as interesting or distracting to a subject as the decisions that they have to make.

When comparing maze types, the blink rate increased for the choice maze over the no choice mazes in most of the subjects. However, blink counts and blink rates did not seem to follow any consistent trend between different difficulty levels for the same type of maze. In particular, there was no majority for whether blink counts increased and decreased from choice to no choice mazes, although many subjects spent less time on the choice maze than on the no choice maze. What was also interesting was noting the vast range in blink rates for different subjects - some individuals had blink rates of one every two seconds, while others did not blink a single time during a half-minute long test.

6.2 Next Steps

Currently, an updated version of the Pupil Player software is being developed to improve the surface tracking accuracy for the newer version of our maze test forms. Once this is complete, it can be used to process the subject tests that were taken using the new forms. Although there are still many recordings that have yet to be analyzed, the current data analysis on the subjects is already telling of trends and patterns in eye behavior throughout the different types of tests. In the future, we can extend this research and check whether these eye behavior patterns are consistent over a larger population. Hopefully we will also eventually apply this study to patients with Alzheimer's and other neurological disorders and evaluate the similarities and differences in their performances. My research will serve as a toolkit of analysis methods to help further the research in this topic.

Appendix A

Additional Documents

Drawing study advertisement:

Subjects Needed for Eye Tracking Study -- $20 for 30 minutes

We're studying hand-eye coordination in a set of simple drawing tasks. You'll be asked to find a path through a maze, draw a common object, or do a simple symbol-digit substitution task, while wearing an eye tracker. The eye tracker is a lightweight eyeglass frame with very small cameras in place of lenses; it weighs only 34 grams (about an ounce).

The study will be done in CSAIL (the Stata building) and will take about 30 minutes. Upon completion you will receive $20 in the form of an Amazon gift card.

Qualifications: Age 18 and over, in good health, able to sit at a desk comfortably for 30 minutes, no restricted head or hand range of motion.

For more information or to volunteer, email drawingstudy@csail.mit.edu.

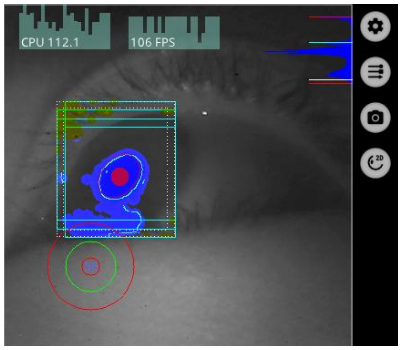
Calibration Instructions

1. Begin by putting the headset on the participant. Place a piece of blank paper on the table in front of them and give them the (capped) pen to hold. Ask them to sit how they would to write. (Note: If the participant owns both glasses and contacts, suggest that they wear their contacts for this test to avoid potential issues with glare.)

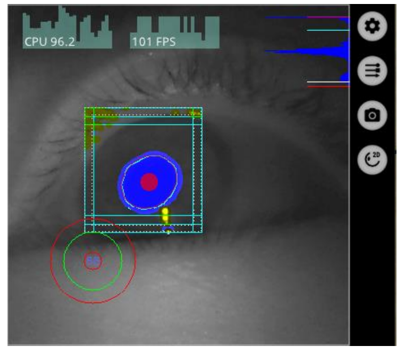
2. Adjust the three cameras:
   - World view camera should see all or most of paper while participant is looking down at it
   - Each eye camera should contain the full range of pupil movement. Ask the participant to look all the way to the left, right, top, and bottom to ensure that the entire range of motion is detected. It's okay if the iris goes out of frame, but the pupil must always be detected.
   - If a pupil is not fully contained in the camera view, gently push the camera around until it can be seen. Ask the participant to close one eye to help you determine which camera corresponds to which eye. Sometimes one eye camera will be flipped, so ask the participant to look up to see if any of the eye videos show the eye looking down instead. Select the option "Flip image display" for that flipped eye.

3. In the settings of each pupil camera, enter algorithm mode and adjust algorithm settings for the pupil detection, if necessary.
   - Adjust minimum and maximum radius (units in pixels) to bound the size of the pupil
   - Adjust pupil intensity range (around 20 is good) so that the blue circle closely follows the circumference of the pupil. We want to avoid the algorithm reading shadows in the eye as potential pupil locations. If you see large blue areas in the shadows of the eye, decrease the intensity range.

- Return to camera mode after adjustments are made.



Intensity range too high



After intensity range lowered

4. Place the calibration target in front of the participant such that it is also in the center of the world view camera. Ask the participant to look at the center of the circle (at the cross) without moving their head.

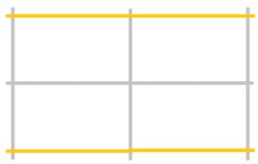
5. Hit “R” to begin the recording. Hit “C” to start calibration. While watching from the screen, move the calibration target (attached to a stick) around in the following ways (visualized below). Ask the participant to follow only with their eyes.

   - Up and down

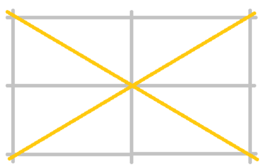
   - Left and right

- Up and down at the leftmost and rightmost sides

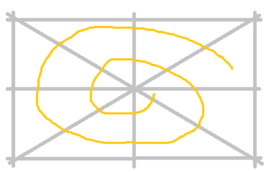
- Left and right at the topmost and bottommost sides



- Across each diagonal



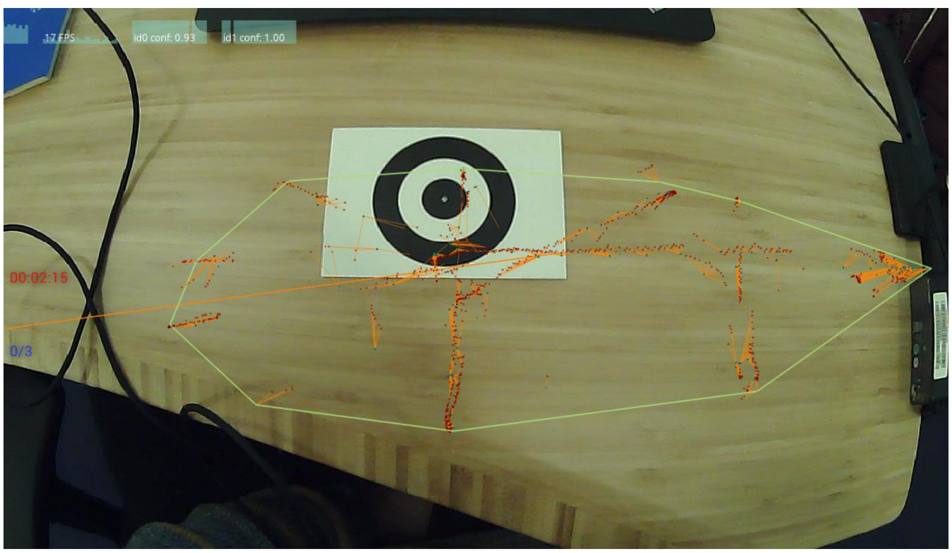
- A spiraling circle



6. Hit “C” again to end the calibration process. Afterwards, a calibration window will appear on the screen. The goal is make this window as large as possible (reaching as close to all edges and corners as possible). This increases the accuracy of the calibration.

   Note: If the calibration comes out poorly, you can go back to Step 4 to redo the calibration process. This can be repeated as many times as necessary.

Good calibration window



Poor calibration window (too small and narrow)

7. Now you can give the participant their test (Maze or SD).

8. After the test is over, hit “R” again to end the recording.

Maze and SD Testing

Preliminaries

Reserve a conference room for administering the test (we need a quiet room without distractions). Conference rooms are in short supply especially during afternoon hours, so plan ahead – match up subject's schedules with available room times. The conference room reservation system is at https://calendar.csail.mit.edu/mrbs; you'll need a CSAIL account to be able to use it.

When you plan an appointment with someone, make sure to get their contact info. That way you'll be able to remind them about the appointment the night before, and can get in touch if they seem to be lost in our confusing building. Also ask whether they routinely use glasses to read; if so they should be sure to bring those to the test.

Have the testing room set up with all necessary equipment ready to go (laptop, light meter, glasses, calibration pattern/stick, a way to take notes if needed, a pen, subject log (attached), etc.)

When they arrive, welcome them, stash their coat, etc., out of the way and ask if they are ready, or need to use the restroom first (to minimize interruptions). Seat them at the table so that they are facing away from any window (even if an interior window), to cut down on possible distractions.

Explain the test in this way, trying to stick fairly closely to this script so that we treat all subjects the same:

We're studying hand-eye coordination in a set of simple drawing tasks. You'll be asked to find a path through a maze and do a simple symbol-digit substitution task, while wearing an eye tracker [point to it]. *The eye tracker is a lightweight eyeglass frame with very small cameras in place of lenses; it weighs only about one ounce. It allows us to record where you are looking from moment to moment.* [Put on the glasses yourself so they can see how they look in place.]

You'll also be using this pen [pick it up, mimic drawing with it], *which works as a regular ballpoint pen but also records your drawing motions digitally.* [Take off eye tracker]

We'll start by having you put on the eye tracker, and we'll adjust it to fit your head. Then we'll ask you sit still while you follow this target [show them] *with your eyes; this make the eye tracker more accurate.*

We'll then ask you to do two simple tests – a maze and a symbol substitution test – wearing the tracker and using the pen. The entire process will take no more than about 30 minutes, and upon completion you'll receive an Amazon gift card worth $20.

Is this clear?

One important thing before we start is for you to read this informed consent form, which describes what you'll be doing and ensures that you understand the goal of this work. Take as long as you need to read and understand it. [Wait until they're done, then ask whether they have any questions about it. Ask them to summarize it in their own words and correct any misunderstandings. Then say:], *please indicate your understanding by signing on the last page.* [Remember to counter-sign it and have an extra copy of the (blank) form to give them.]

Testing

Have them comfortably seated in the appropriate chair. Make sure subject is wearing theit glasses if needed. Put on the tracker and make appropriate adjustments. Then say:

> *Now I'd like you to sit fairly still and follow this target with <u>just your eyes</u>, without moving your head.*

Instructions for administering both the maze and SD test are attached. Use the wording suggested there so that everyone receives the same instructions.

After they are done, make sure to carefully collect the tracker and pen, log the test date time, and the fact that they got an Amazon card. Ask if they have any questions or concerns related to participating in this study, and point out to them that there is contact information for Professor Davis on page 2 of the informed consent form that they've been given a copy of. Give them the Amazon card and thank them for their interest and participation. If they say they're interested in the work, tell them we'll let them know when we have results analyzed and written up.

Digital Maze Test

General Instructions

- Present the no-choice maze (the right side of the form) first.
- Give the digitizing pen to the examinee AFTER the instructions and just before they are to write on the paper.

Administering the test

- When ready to test, check the pen "start" box (at top right of the form) just before presenting the paper to the examinee.
- Show the examinee the paper and say: *"This page has two parts. The top part is a large maze* (point to the maze) *and the bottom part is small practice maze"* (point to the sample maze*).*

  The examiner should partially cover the large maze by lightly resting their hand over the maze while pointing to the small maze and giving the instruction, and while the subject completes the sample maze.
- *"On the bottom part* (point to the sample maze), *you have a practice maze with a single path. The beginning of the path has a green arrow and the end has a red arrow. I'd like you to start by putting your pen on the green arrow and then draw a line through the path as quickly as you can and without bumping into any walls, stopping at the red arrow. Any questions? Okay, go ahead."*
- After the examinee completes the sample say, *"Good. Now I'd like you to try the bigger maze up top"* (point to the bigger maze)

- *"As before, you should start by putting your pen on the green arrow. When I tell you to start, you should draw a line through the maze as quickly as you can, without bumping into or cutting through any walls. If you make a mistake, keep your pen on the page and just backtrack. Do you understand?"*
- If there are questions, repeat the instructions or relevant portions of the instructions as many times as necessary.
- *"Okay, go ahead."*
- When the examinee completes the right-side maze, turn the paper over so the left-side maze is on top, and give the same instructions as you did for the right-side maze (there is no maze trial with the left side maze).
- When the examinee has completed both mazes, use the digital pen to write subject identification information at the bottom .

Testing Issues

- The examiner should watch the examinee carefully. If the examinee:
  - lifts the pen say, "*remember to keep your pen on the page*"
  - bumps into a wall say, "*try not to bump into the walls*"
  - cuts through a wall, point to the path where they cut though the wall and say, "*put your pen here and remember not to cut through the walls*"
  - lifts the pen and moves it to another location, point to the place where the pen was lifted and say, "*keep your pen on the page and backtrack*"
  - has difficulty seeing the pen point because their hand is in the way, say "*try your best to keep the pen on the page the entire time"*

Administering the digital Symbol Digit Test

General Instructions:

- Present the Symbol-Digit subtest (the top half of the form) first. Explain the test using the script below. When you are ready to proceed, put a checkmark in "start box" in the upper right hand of the form just before handing the subject the pen.
- The examinee completes/writes all of the sample items.
- When complete, the examiner takes the pen, **writes necessary identification in the upper left area of the test**, and then makes a checkmark in the upper left box:

Symbol-Digit Trial

Explain the first half of the test by saying:

This part of the test has two parts. The top part is an answer key (point to the answer key) *and the bottom part is where you write your answers* (point to the response section*).*

Up top at the answer key (point to answer key) *you have boxes that have pairs of shapes and numbers. The top boxes all have shapes and the bottom boxes all have numbers. Each shape has its own unique number.*

On the bottom part (point in general direction of the response section), *you have top boxes with shapes, but the bottom boxes are empty. I want you to use the answer key above to write in the bottom boxes the number that goes with each shape.*

EXAMINER DOES NOT WRITE IN ANY SAMPLE ITEMS

Look at the first shape. (point to first shape in the Sample section) *You see that it's a cross. You look up to the answer key to find the cross and you see that it is paired with the number one. You put the number one in the empty box where it belongs.* (point to the empty square below the cross)

Now look at the second shape and see that it's a triangle. Look back up to the answer key, find that the triangle is paired with the number 10 and then put the 10 in the empty box." (point to the empty box below the triangle)

You should point to the answer key and repeat the instructions as many times as necessary. Proceed to the sample when the examinee understands the task.

Give the examinee the digital pen.

Let's do the sample. Start at the beginning and take them in order, stopping at the bold line here. Remember take them in order – no skipping. (point to the bold line that divides the sample items from the test items)

Watch as the examinee completes the sample. You should point out any mistakes that the examinee makes at the time of the error and ask them to correct their error.

After the sample is completed say:

Good. For the next part BOTH speed and accuracy are important. When I tell you to 'Go' you should start here… (point to the first test item – the cross) *…and continue on to the end at the stop sign.* (at the same time point generally to the second and fourth lines using a sweeping z-shaped gesture – stopping near the stop sign)

Remember to work as <u>quickly and as accurately</u> as you can – without skipping any items.
Are you ready? Okay, GO.

When the symbol digit subtest is completed, immediately turn over the page and expose the Digit Copy subtest.

<u>Digit-Copy Trial:</u>

This page is a little different, but it also has an answer key (point to the answer key) *and a place to write your answers* (point to the response section*).*

Up top at the answer key… (point to answer key) *…you now have boxes that have pairs of numbers. The top boxes and the bottom boxes <u>each</u> have the <u>same</u> number. <u>The numbers in each pair are the same</u>.*

On the bottom part … (point in general direction of the response section), *you have top boxes with numbers, but the bottom boxes are empty. I want you to write in the empty boxes the numbers that match the ones in the top boxes – that is, you should copy the numbers in the empty boxes below them.*

Let's do the sample. Start at the beginning and take them in order, stopping at the bold line here. Remember take them in order – no skipping. (point to the bold line that divides the sample items from the test items)

Watch as the examinee completes the sample. You <u>should</u> point out any mistakes that the examinee makes at the time of the error and ask them to correct their error.

After the sample is completed say:

Good. For the next part BOTH speed and accuracy are important. When I tell you to 'Go' you should start here… (point to the first test item – the number 1) *…and continue on to the end at the stop sign.* (at the same time point generally to the second and fourth lines using a sweeping z-shaped gesture – stopping near the stop sign)

Remember to work as <u>quickly and as accurately</u> as you can – without skipping any items. Are you ready? Okay, GO.

When the test is complete immediately proceed to the Symbol-Digit recall trial.

<u>Symbol-Digit Paired-Recall trial:</u>

Do you see here that you once again have some shapes, and below them are empty boxes? I'd like you write in the empty boxes the numbers that belong there, just as you did a few minutes ago. Try to remember which number matches the shape. If you're not sure, try to guess.

SUBJECT LOG

Our ID JW001 / SL001	Name (first, last)	Date/Time	Room illum. (Lux)	Received Amazon card

Bibliography

[1] Kiyohiko Abe, Hironobu Sato, Shogo Matsuno, Shoichi Ohi, and Minoru Ohyama. Automatic classification of eye blink types using a frame-splitting method, 2013.

[2] Simone Benedetto, Marco Pedrotti, Luca Minin, Thierry Baccino, Alessandra Re, and Roberto Montanari. Driver workload and eye blink duration, May 2011. Retrieved from https://doi.org/10.1016/j.trf.2010.12.001.

[3] Pieter Blignaut. Fixation identification: The optimum threshold for a dispersion algorithm, 2009. Retrieved from https://link.springer.com/content/pdf/10.3758

[4] Melissa Patricia Coral. Analyzing cognitive workload through eye-related measurements: A meta-analysis, 2016.

[5] Bryn Farnsworth. 10 most used eye tracking metrics and terms, August 2018. Retrieved from https://imotions.com/blog/7-terms-metrics-eye-tracking/.

[6] John Findlay and Robin Walker. Human saccadic eye movements, 2012.

[7] Matthias Hartmann and Martin H. Fischer. Pupillometry: The eyes shed fresh light on the mind, February 2014. Retrieved from http://dx.doi.org/10.1016/j.cub.2014.02.028.

[8] Ian P. Howard. Binocular vision and stereopsis, 1995.

[9] Moritz Philipp Kassner and William Rhoades Patera. Pupil : constructing the space of visual attention, 2012. Retrieved from http://hdl.handle.net/1721.1/72626.

[10] The parameters of eye movement. Retrieved from https://www.liverpool.ac.uk/ pcknox/teaching/Eymovs/params.htm.

[11] Farah Maqsood. Effects of varying light conditions and refractive error on pupil size, 2017. Retrieved from https://pdfs.semanticscholar.org/2a8d/484acb83c0c179ffbe0681afb06d16a7cb17.pdf.

[12] Sebastiaan Mathot, Jasper Fabius, Elle Van Heusden, and Stefan Van der Stigchel. Safe and sensible preprocessing and baseline correction of pupil-size data, January 2018. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5809553/.

[13] National Institute on Aging. Alzheimer's disease fact sheet, August 2016. Retrieved from https://www.nia.nih.gov/health/alzheimers-disease-fact-sheet.

[14] Jami Pekkanen and Otto Lappi. A new and general approach to signal denoising and eye movement classification based on segmented linear regression, 2017. Retrieved from https://www.nature.com/articles/s41598-017-17983-x.

[15] Dana Penney, November 2018. Conversation.

[16] Marta LG Freitas Pereira, Marina von Zuben A Camargo, Ivan Aprahamian, and Orestes V Forlenza. Eye movement analysis and cognitive processing: detecting indicators of conversion to alzheimers disease, July 2014. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4096446/.

[17] Harvey Richard Schiffman. Sensation and perception. an integrated approach, 2001. Referenced from https://bionumbers.hms.harvard.edu/bionumber.aspx.

[18] Ralf van der Lans, Michel Wedel, and Rik Pieters. Defining eye-fixation sequences across individuals and tasks: the binocular-individual threshold (bit) algorithm, November 2011. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3048294.