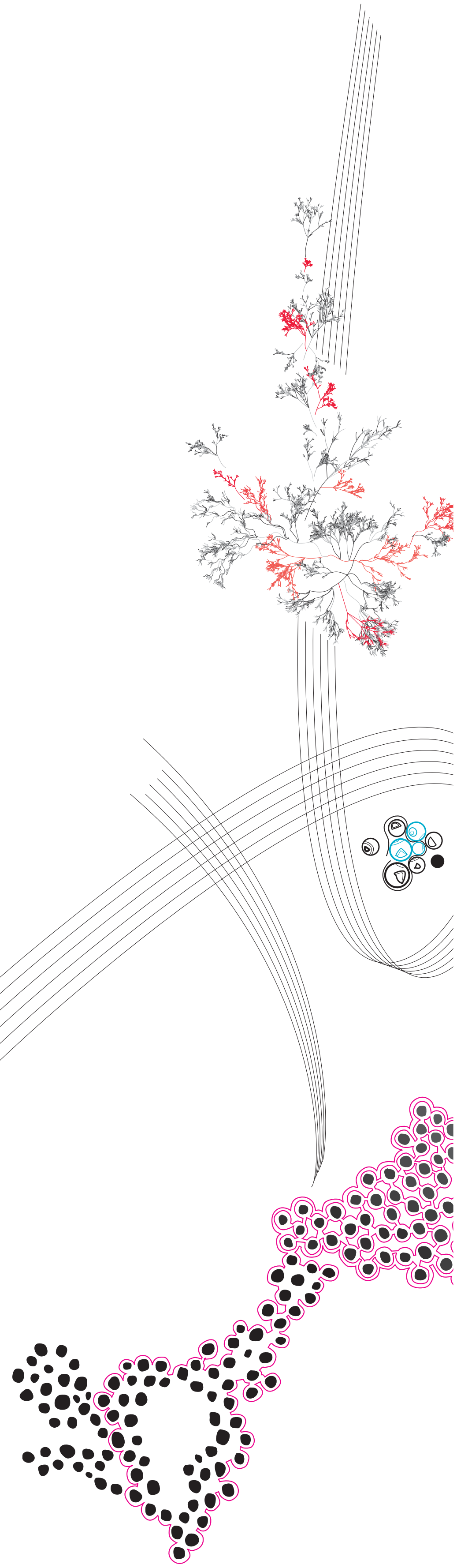


BSc Thesis Applied Mathematics

Modeling an eye gaze training task for people with PIMD

P.D. de Regt

Supervisor: J. Goseling, D. Reidsma

July, 2021

Department of Applied Mathematics
Faculty of Electrical Engineering,
Mathematics and Computer Science

UNIVERSITY OF TWENTE.

Preface

This article was written as a bachelor's thesis at the faculty of Electrical Engineering, Mathematics and Computer Science at the University of Twente, as a closing assignment for the double degree Technical Computer Science and Applied Mathematics over the period of three months, April, March and June, in 2021.

The research was inspired by an initiative from Koninklijke Visio Amsterdam, an organisation that offers information and support for people with a visual impairment, including people with other impairments or disabilities next to a visual impairment. I want to thank them for the idea. In particular, I want to thank Wieneke Huls and Jan Koopman. Jan, who helped me with some of the technical aspects, was always very ambitious, positive, and eager to help, and created tasks that have laid the foundation on which some parts of the research are built. Wieneke's extensive knowledge of and experience with people with PIMD, which she was so kind to share with me, was very valuable, and gave insights that could not have been achieved with only literature. She always was very supportive and open for any questions I had. I want to thank them both for their time and help, which was extremely valuable. Without it, this research would not have been possible. I hope this research can be of some value for them too.

Next, I want to thank my friends, parents and brother. They were always there for me if I was feeling down, to cheer me up or offer me useful advice. If it was not for them, I would not have overcome the stressful and mentally tough times that I faced. Each in their own way, they supported me and therefore indirectly made their contribution to this article.

Finally, I want to thank my supervisors, Jasper Goseling and Dennis Reidsma. First of all, I want to thank them for their enthusiasm, which showed when they agreed to supervise this research, which I had personally proposed, and during the meetings, where they always gave very useful feedback and attempted to inspire me to improve the research. Even if not everything went as was hoped, they remained supportive and willing to provide me with feedback and advice. Without them, I would not have been able to perform this research.

Hopefully you enjoy reading this thesis!

Modeling an eye gaze training task for people with PIMD

P.D. de Regt

July, 2021

Abstract

This research studies an eye gaze training task that is used with people with Profound Intellectual and Multiple Disabilities (PIMD) to help them in learning to use an eye tracking device. These individuals have difficulty communicating, but eye tracking technology, together with a communication computer, can serve a form of Augmentative and Alternative Communication (AAC) to overcome the physical barrier of communication that is faced. By improving the training process, more people with PIMD can profit from this technology. For this research, data coming from an eye tracker is studied and characterized. Then, a training task is defined and formulated. Using these findings, a method is proposed to evaluate user performance in the given task, based on accuracy and fixation duration. This result can help professionals involved in the training to judge the progression of the user, and can be used to improve the task by developing a method which uses Dynamic Difficulty Adjustment (DDA) to adapt the task to the individual and make it more effective.

Keywords: eye gaze, AAC, mathematical model, PIMD

Contents

1 Introduction

Communication might be one of the most important abilities of humans, and is also vital for participating in society. However, being able to communicate properly is not straightforward for everyone. Some individuals have much more difficulty to express themselves in a way that is understandable for others. Some individuals also find it hard to understand what others are trying to communicate. One such a group that suffers from both these challenges is the group of individuals with Profound Intellectual and Multiple Disabilities (PIMD). They often lack the ability to communicate through traditional means, such as speech or writing.

To help individuals with PIMD overcome this problem, eye gaze devices can be used to help them communicate using their eyes. The eye gaze device will mainly help to overcome the physical aspect of communicating. To help individuals become familiar with the eye gaze device, and improve their ability to use the technology, training can be used. This training can come in the form of simple games or tasks to be played or performed by the individual. Next to helping the individual get used to the eye gaze device, data from the training can also be used to say something about how well the individual is able to use the eye gaze device.

However, this process is not without its flaws. It can often be difficult to determine how well a person is fit for the technology. Also, the training tasks have a set, static difficulty. This, amongst other problems, can lead to some individuals being excluded from using eye gaze technology, as it might appear that they do not have the ability to use the device. However, if this is indeed caused by the technology, and not the ability of the individual, this exclusion would be unjustified.

This research aims to solve this issue, attempting to look into the training tasks and improve their ability to identify the skill level of the user. For this, a foundation is built by giving relevant background knowledge, the research question and methods are noted and the model is presented and discussed.

2 Background

In this chapter, relevant theory,on which this research is built, is discussed. Important findings from earlier research, coming from different fields, are summarized, and with this, knowledge is established that will be used later on in the research.

2.1 People with Profound Intellectual and Multiple Disabilities

As the research deals with people with Profound Intellectual and Multiple Disabilities (PIMD), it is important to identify some aspects that characterize individuals in this group. People with PIMD, as the name suggests, have a profound intellectual disability, which means they have an estimated intelligence quotient of 20 or lower, and therefore have severely limited understanding [1]. Next to that, they also suffer from a profound motor disability, as well as a number of additional severe or profound secondary disabilities or impairments [2]. This may include impaired hearing, vision or movement, as well as other problems.

Individuals with PIMD often have much difficulty to communicate with their caregiver and others, and can be trapped inside their own body. As they suffer from a profound intellectual disability, combined with physical or sensory impairments, both can be a big obstacle when trying to communicate: they may lack the mental capacity to properly communicate, or their physical impairments prevent them to be able to express themselves properly, or possibly even both [3]. This can be quite a challenge for the caregiver, a position often filled by a family member: Due to their disabilities, individuals with PIMD require a high level of support from others, not only due to the needed medical attention, but help is also needed with daily tasks, such as eating, dressing, using the toilet and many other common efforts. During these tasks, it can be difficult to understand what the individual with PIMD is trying to express, an obstacle that these caregivers face daily.

2.2 Eye-tracking technology for people with PIMD

A technology that could help solve this problem is eye tracking technology. Eye tracking uses infrared light reflected by the eyes to determine the position of the gaze of the user. This is done frequently, such that the many samples can provide a good estimation of the activity of the user's gaze. How frequent depends on the eye tracker, but the sampling frequency does have an impact on the accuracy of measurements [4]. Eye trackers can come in multiple forms: as a remote desktop system, but also as head-mounted desktop systems and mobile eye tracking systems [5]. However, head-mounted desktop and mobile eye tracking systems are intrusive, while a remote desktop system is not. A remote desktop system is also easily movable, so it can be used at numerous locations. Therefore, this is most appropriate when dealing with people with PIMD, as both physical contact and a remote location can be big obstacles and reasons not to use the technology. However, some limitations of these eye trackers are that head movement could negatively impact the data quality and that the eye tracker data is mostly valid when the participant is looking at the monitor that is part of the remote desktop system.

The technology can be used for two main purposes. The first is that the eye tracker data can say something about the user. Using the data, the gaze of the user can be studied, to see for example what grabbed their attention or how well the user was able to focus on a certain object. The second purpose is that the user can try to actively use the eye tracker to interact with the computer. In this way, the user can use its eyes to control the computer

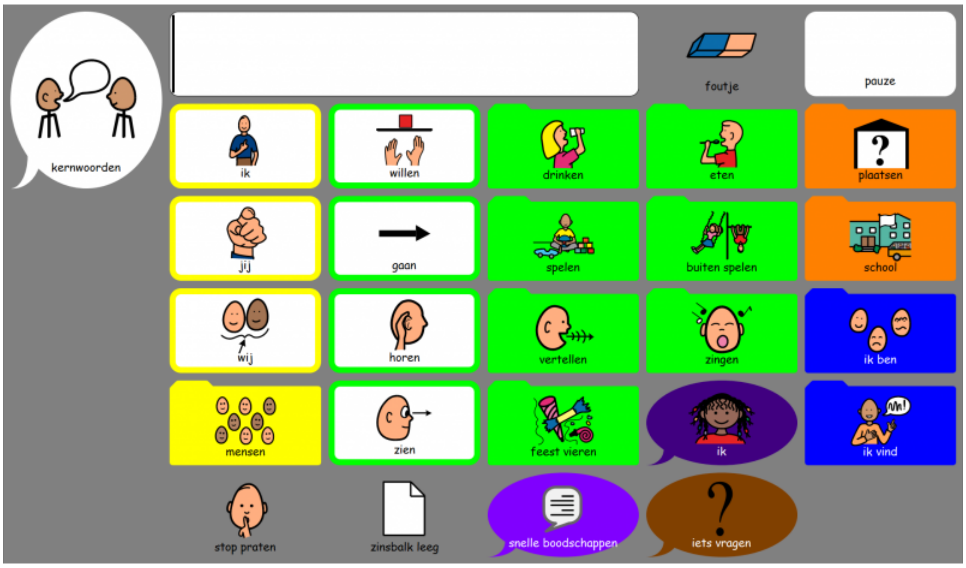


FIGURE 1: ALOHA: an example of a Dutch core vocabulary

and for example use it to communicate. When communicating is the goal of using eye tracking technology, the device is then used as a form of Augmentative and Alternative Communication (AAC): People with PIMD will be able to use it to communicate with others better, more easily and/or more specifically. This idea was shown to be promising for individuals with Rett Syndrome [6] [7], which increased the interest in using the technology as a form of AAC.

The way eye tracking is used as AAC is by using the user's gaze to control a communication computer, a device running certain software to allow the user to express themselves by selecting symbols or words from a given vocabulary, or possibly perform other actions such as accessing the internet[1]. The vocabulary can take on many forms, and can also be personalized by adding custom words and symbols and changing the size and amount of options, although this is not often done in practice. More common is the use of a core vocabulary. One example of this is ALOHA[2], seen in figure 1.

For people with PIMD, using eye tracking technology can be quite a challenge: not only do they need to understand what the device does and what they can do with it, namely perform actions on the computer, but they should also learn how to actually use the device. Koninklijke Visio approaches this problem by using simple eye gaze games, or tasks, as training for individuals with PIMD, in the project "Sprekende Ogen 2.0" [8]. Eye gaze games offer a great opportunity to both help the user become familiar with the eye tracking device, thereby improving their capability to use the device, and collect gaze data about the user. Using games to train or educate has become relatively common, and can help increase interest and learning in the subject [9]. However, there is no extensive background on using training tasks for the specific case of teaching people with PIMD to work with an eye gaze device.

To help individuals with PIMD understand what the eye tracking device can do, the training by Koninklijke Visio is split into multiple phases, based on the Eye Gaze Software Curve[3] in figure 2. The first phase tries to teach the user to look at the screen. Phase

---

[1] An example is the Tobii communicator 5 software (https://www.tobiidynavox.com/software/windows-software/communicator-5/)

[2] https://rdgkompagne.nl/communicatiehulpmiddelen/vocabulaires/aloha-voor-mind-express/

[3] https://www.callscotland.org.uk/downloads/posters-and-leaflets/eye-gaze-software-curve/

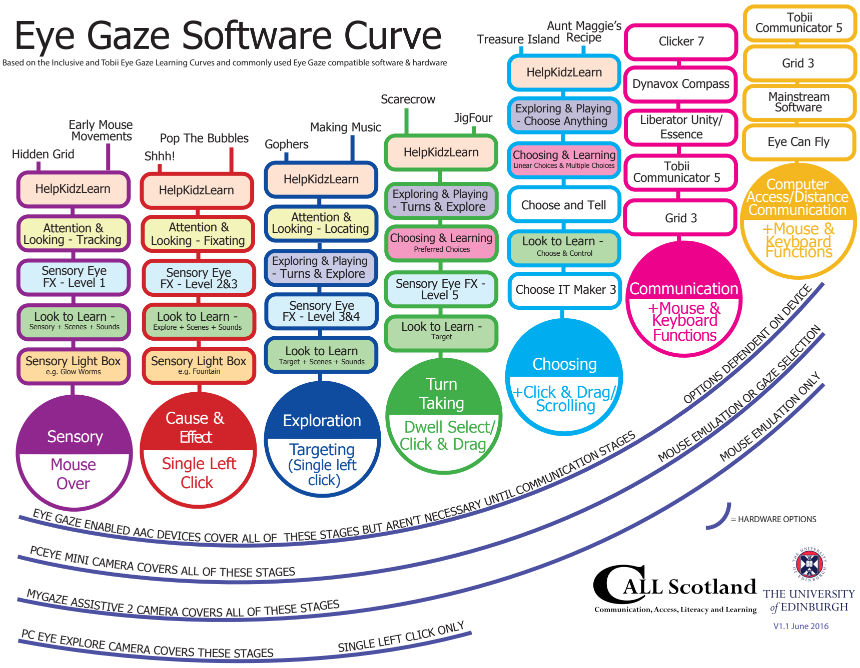


Figure 2: Eye Gaze Software Curve, published by CALL Scotland

2 learns the user that looking at a specific object on the screen can cause something to happen. In phase 3, the user learns that there may be several options to explore and select. By going through these phases step by step, the user gradually learns what the eye tracker does, and how it can be used, thus helping them overcome the possible mental challenge of using eye tracking technology.

2.3 Eye movement

To interpret data coming from the eye tracking technology, it is importnat to know how the human gaze works. In general, the gaze is said to be composed of two main events: fixations and saccades.

Fixations are periods of time in which the gaze of the eye dwells on a certain stimulus or area, with no substantial movement. If a person fixates on a certain stimulus, this will most likely mean that their attention is focused on this stimulus. Duration of a fixation or number of fixations can be used to infer something about the amount of attention a person gives to a certain area or stimulus. A fixation with a duration as short as 100ms can already be of significance when attempting to extract information about the eye movement of a person [5].

Saccades are defined as the rapid shift of eye gaze that happen between fixations. During these saccades, no new information is obtained, as the eye movement is so fast that only a blur would be perceived [10]. Important properties of these saccades that can be used for research are velocity and direction.

There are other types of eye movement, namely pursuit, vergence and verstibular eye movement. Pursuit is following a moving target with your eyes. Vergence is the rotation of the eye to fixate on a nearby object. Vestibular eye movement is eye rotation to compensate for head and movement and maintain the same direction of vision [10].

During a fixation, the eyes do not completely stand still. Small eye movements still occur during these periods of fixation. These eye movements can be categorized as nystagmus, drifts and microsaccades. For some purposes however, these movements can be considered as noise [10].

To extract the mentioned measures from eye tracker data, an event detection algorithm can be used, such as the one developed by Nyström and Holmqvist [11]. This algorithm can detect fixations, saccades, and glissades, wobbling movements at the end of saccades. With the information about all these events, something can be said about the gaze and attention of the user of the eye tracker.

2.4 Dynamic Difficulty Adjustment

For all games, so also video games, difficulty is an important factor for the enjoyability, and in the case of training games or tasks effectiveness, of the game. With a static difficulty, the user may find the game too difficult or too easy. To prevent this, the technique of Dynamic Difficulty Adjustment (DDA) can be used, which aims to keep the user interested and to offer a satisfactory challenge level, by adjusting scenarios, parameters and behaviors in games [12].

A common approach to model the player investment in the game is by using the flow model, developed by M. Csikszentmihalyi [13]. The goal is to ensure the player is kept in the flow channel: away from states where the game is too easy or too difficult [14]. This concept may also be useful in the training tasks for using an eye tracker.

3 Research questions and methods

For this research, the goal is to improve eye gaze training tasks for people with PIMD. The focus lays on the physical ability to use the eye gaze device, thus for now ignoring the possible mental challenges of understanding what the device does. This issue is assumed to be solved by the use of different phases in the training and in this research only a single phase will be considered. Also, the ability to use the eye tracking device will be operationalized as the ability to control eye movement. The question this research is trying to answer is the following. How can tasks performed using an eye tracking device help determine the ability of Profound Intellectual and Multiple Disabled (PIMD) people to control their eye movements?

The proposed solution consists of the following process. The user will perform a certain task with the eye tracker. During this task, the eye tracker will collect data about the gaze of the user. This data will be used to determine an indication of the performance of the user. A possibility is then to use this indication to adapt the difficulty of the task. The process is also visualized in a flowchart in figure 3.

For the process to work correctly, the different parts have to be looked at and specified precisely. Therefore, the following sub-questions are defined to elaborate on the process, and together answer the main research question.

First, the data coming from the eye tracker is analyzed. To make the data usable, knowing what the data looks like is vital. Possible issues should be uncovered, and other details about the data should be discovered, such as the amount of variance and noise. This is the goal of the first sub-question - What are the different properties, flaws and behavior of the data coming from the eye tracker? This question is answered by a small exploratory experiment, where data will be collected using the eye tracker, the results will be analyzed and conclusions will be drawn about how the data can be interpreted and used.

Second, the task to be performed by the user should be considered. The following sub-question is answered: How can an eye gaze training task be characterized? This is done by first describing the chosen task, and then formulating this task in an abstract way, such that the data fits into this model of the task. Finally, the aspects of the task determining its difficulty are discussed.

Finally, the data should be used to determine an indicator for the performance of the user of the eye tracker concerning the task. This model will answer the second sub-question: How can the data from the eye tracker be used to determine the performance of the user? The answer is given by specifying a model which translates the data coming from the eye tracker into a value that gives an indication of the performance of the user.

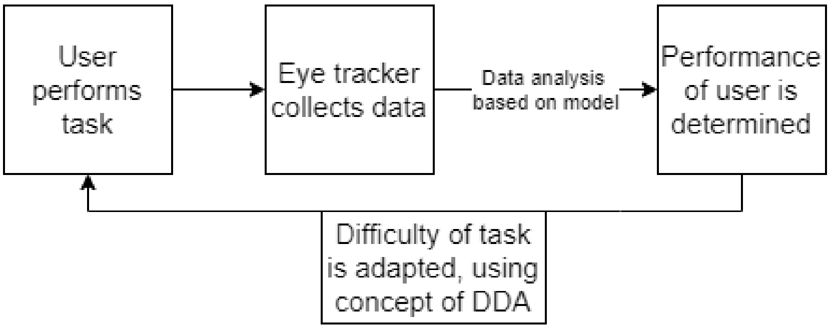


FIGURE 3: Flowchart visualizing the process

4 Data quality

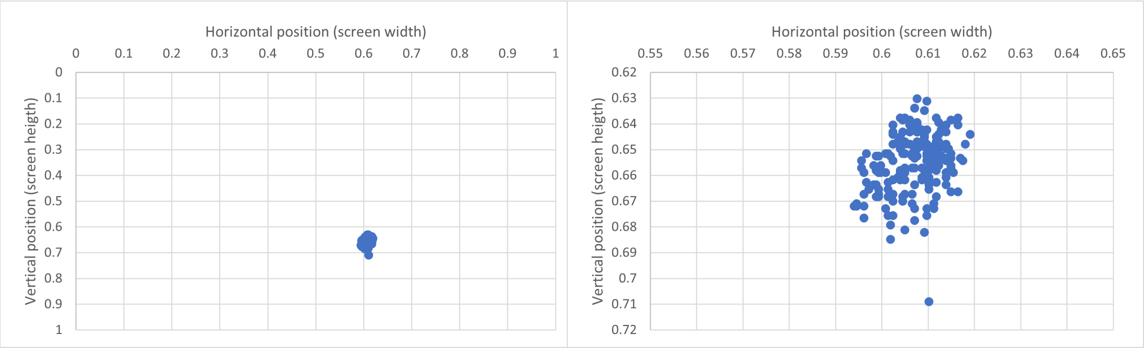
To construct a model which uses data to estimate performance in an eye gaze task, first the data itself should be acquired and the data should be analysed for possible flaws, noise and issues. The information and conclusions drawn from this will have implications for the model of the task and the evaluation of user performance. The chapter thus investigates the block "Eye tracker collects data" of the flowchart in figure 3.

4.1 Method

For this exploratory research a single participant was used, namely myself. As I do not have any relevant visual or other physical impairments, the data should reflect relatively accurately the position of my gaze. The task to be performed during data collection was to look at specific points on the screen, namely the point in the center of the screen, the points at the middle of the edges of the screen, and the points in between those edges and the center of the screen. These points were specifically marked, so the point to be looked at was unambiguous.

For the data collection, the eye tracker was set up in a way such that it could capture the eyes of the user properly, and the background, so behind the user, was relatively dark, as much light behind the user can harm the data quality. Then, the eye tracker was activated and the data collection could start. Due to technical constraints, the data was collected by attaching the gaze position, as estimated by the eye tracker, to the cursor and reading out the coordinates of the cursor. A sample was taken every 0.1 seconds, and the samples were taken over a period of 20 seconds for each point on the screen.

4.2 Results



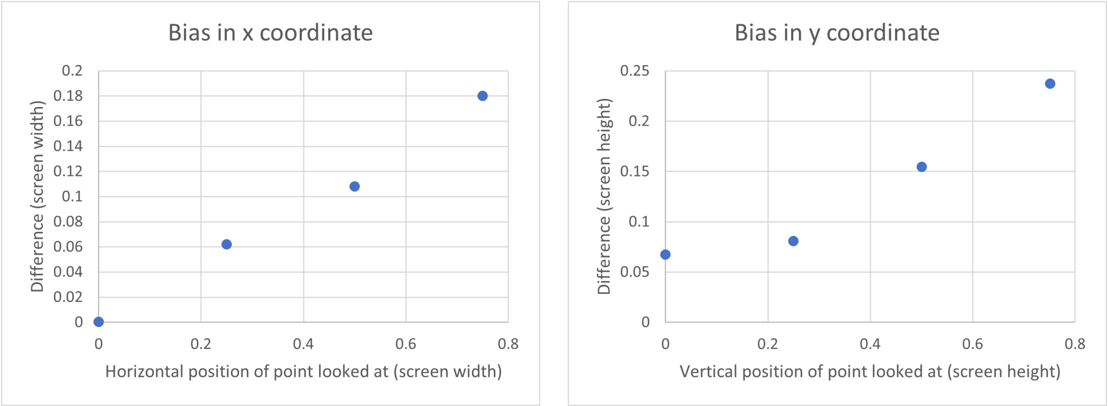
(A) Relative to the entire screen

(B) More locally specific

FIGURE 4: Scatter plots of measured gaze positions when looking at the center of the screen

The measurements taken during the task of looking at the center of the screen can be seen in figure 4. Both horizontal and vertical position of the measured gaze samples are normalized to screen width and height respectively. Also, the top left of the screen is considered as the origin, and the vertical position is inverted, as the coordinates of the cursor also behave this way.

One thing to note is that the mean of the data seems to be quite far off of the actual center of the screen, which is the point (0.5, 0.5): none of the samples are close to this point. There seems to be a bias in the data. In figure 5, the difference between the position looked at and the mean of the measured gaze samples is plotted, for both horizontal and vertical position. For both graphs, the other dimension is constant at a value of 0.5.

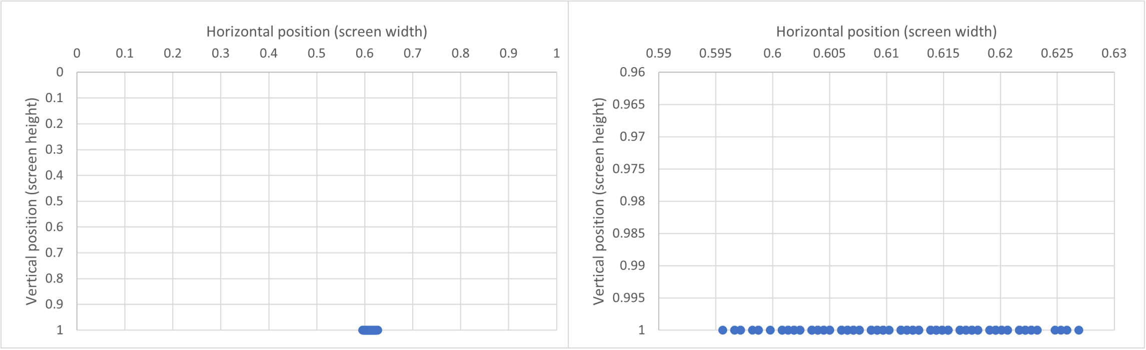


(A) Difference in x on varying horizontal positions (vertical position constant and centered)

(B) Difference in y on varying vertical positions (horizontal position constant and centered)

FIGURE 5: Difference of position looked at and measured coordinates

Figures 6 and 7 are plotted to study how data behaves at the edge of the screen, where the figures represent data collected while looking at the center of the bottom edge and the center of the right edge respectively. In both cases, the coordinates are constant in one dimension (this is partially caused by the bias), but still vary in the other dimension.



(A) Relative to the entire screen

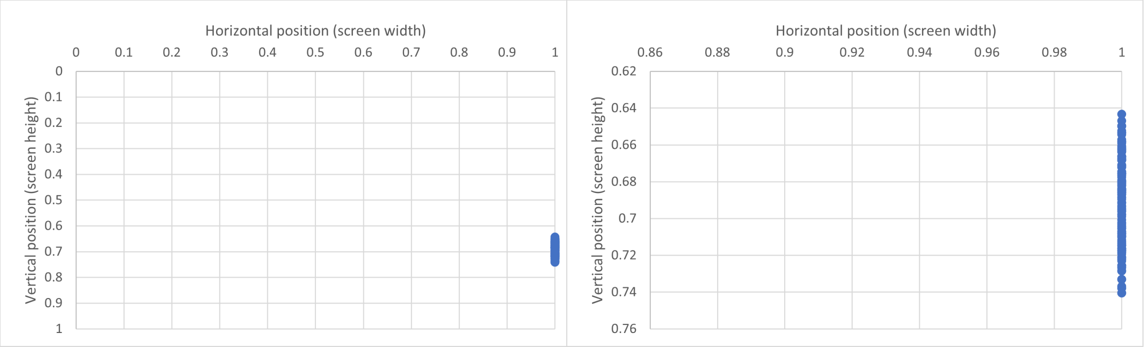
(B) More locally specific

FIGURE 6: Scatter plots of measured gaze positions when looking at the center of the bottom edge of the screen

A final observation is that, when the eye tracker cannot detect the gaze of the user, the position of the cursor will not be updated, so the cursor will remain idle. When the gaze is captured, the cursor will be updated. As seen in the figures 4, 6 and 7, the detected gaze will vary. The data shows that two consequent samples with identical coordinates are very unlikely, although they are more likely to happen at the edges of the screen, due to one dimension being constant.

4.3 Conclusion and implication for the model

The results show that the amount of bias is significant and not constant, but rather linearly dependant on the relevant coordinate position, as can be observed in figure 5. Although the presence and amount of bias may depend on other factors, such as environment and



(A) Relative to the entire screen (B) More locally specific

FIGURE 7: Scatter plots of measured gaze positions when looking at the center of the right edge of the screen

set up, and the bias may heavily differ for other eye trackers, it should certainly be taken into account in this research. One implication of this bias is that this makes it difficult to collect valid data for points laying in certain regions of the screen which are projected on the side of the screen by the bias. For this study, these regions should consequently be avoided where possible and no target objects should be placed here.

Knowing if the gaze is detected by the eye tracker is also relevant for the model. As observed in the results, identical consequent samples are unlikely to occur when the gaze is detected. This is a consequence of both micro-saccades in the gaze of the user and noise from the eye tracker. Hence, the idleness of the cursor can function as an indicator of whether the gaze is detected by the eye tracker, where 5 identical and succeeding samples can be interpreted as the absence of the gaze. This threshold of 5 samples is chosen to ensure the gaze is absent, and will in this case span over 0.5 seconds, as a sampling frequency of 10 Hz was used for this experiment. A different sampling frequency could require a different threshold.

Finally, the behavior of data when the measured gaze is at the edge of the screen is considered. The exact position of the gaze cannot be determined if it falls outside of the screen, but the presence of the gaze can still be detected if the measured gaze is either vertically or horizontally still in line with the screen. While one coordinate is constant, the other coordinate varying indicates the presence of the gaze. This can be seen in figures 6 and 7, where we see that the coordinate still in line with the screen still shows variance. Because the presence or absence of a gaze still can make a difference in the model, this implies that the model should also take measurements into account that are projected on the edges of the screen, if one dimension still shows variance.

5 The eye gaze task

In this section, the task that is used for this research is described, and formulated mathematically. It thus covers the 'User performs task' box from the flowchart of figure 3. This step is vital to be able to later define mathematically an indicator for the performance of the user, as it can be based on the environment which is drawn up in this chapter. First, the task will be described like it is in reality. Then, it is translated to an abstract representation of the task. Finally, a number of aspects determining the difficulty of the task are mentioned and a choice is made which ones are most suitable for this research.

5.1 The task in reality

For this model, a task is chosen which falls under phase 2 in the Eye Gaze Software Curve, as introduced in the background and shown in figure 2. This means that the task intends to teach the user cause & effect. The task that is used is based on a task used by Koninklijke Visio in their training for people with PIMD, and is to be performed on a monitor using an eye tracking device.

The goal of the task is to look at an object, which is fixed. The object could be a picture, a video or an animation, but should have a fixed size. This object will be placed somewhere on the screen, and the user is incentivized to look at this object. This can be done by making it the only interesting thing on the screen, or by providing an auditory or visual stimulus to the user when its gaze is on the object. Fixating on the object for a set length of time will provide the user with a reward, such as an animation or a sound. This not only encourages the user to continue looking at the object and/or to later return their gaze to the object, but also teaches the concept of cause & effect: by looking at a certain thing on the screen, the user made the animation or sound happen.

A possible example of the task described can be seen in figure 8. Some hypothetical eye gaze data is also included in the figure, to give an idea how the gaze data could look in the context of the task. The object is in this case the heart shaped figure, and could light up or spin around, or music could be played as a reward if the user fixates on the object for long enough.

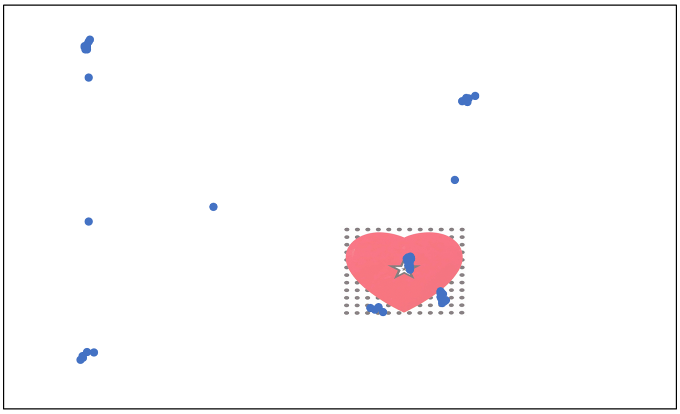


FIGURE 8: Possible version of the described task, including hypothetical eye gaze data

5.2 Mathematical formulation

Not everything is relevant for this research, so to extract only those important aspects, the environment will be simplified: we are mainly interested in the computer screen on which the task is performed, and thus on which the eye movements of the user are registered. The rest of the environment, including other objects or people outside the screen, will, even though it most likely affects the user and its performance, not be taken into account.

As the computer screen is two dimensional, it can be represented as a two-dimensional Cartesian coordinate system. Let the center of the coordinate system be the top left corner of the computer screen. Also, let the y-axis be inverted, such that it starts at 0 and it increases downwards. These choices are made based on the format of the data, covered in chapter 4. Let the width of the computer screen be w and the height be h.

Then there is the object at which the user should look. As this object could take on many shapes, it is simplified in this model to a point in the coordinate system, corresponding to the center of the object, and a radius, which is the maximum distance from the center to any other point included in the object. The center point of the object will be denoted by (x_o, y_o), the radius by r. The object will thus be the set of points

$$O = \{(x, y) \mid |(x - x_o, y - y_o)| \leq r\} \tag{1}$$

with $r \geq x_o \geq w - r,\ r \geq y_o \geq h - r,\ 0 \geq r \geq \frac{1}{2} * \min(w, h)$. The conditions ensure the object is located on the screen and fits on the screen as well.

5.3 Task difficulty

A very important property of the task is its difficulty. This is very important to evaluate the performance of the user as well. There are multiple factors that determine the difficulty of the task.

1. The size of the object
   If the size of the target object is bigger, the task is easier. This size is also represented in the model by the value of the radius r.

2. The time of fixation needed to receive the reward
   The longer the user needs to focus on the object to receive the reward, the harder it is to 'complete' the task successfully.

3. The visual attractiveness of the target
   If the object grabs the attention, this will make the task easier. Examples are making the object more colorful, or using an animation.

4. What the rest of the screen looks like
   This relates to the previous item: the more bland the rest of the screen is, the more the object will stand out, making the task easier. A busy background will make it harder to focus on the target object.

5. The placement of the object
   The position of the target object could also influence the difficulty. A more central object could be easier, but this also depends on the user.

And there are most likely many other aspects that play a role in determining the difficulty of the task. However, not all are easily quantifiable. For this reason, the decision was made to only consider item 1, the size of the object, and item 2, the fixation time, in this research. These will therefore play a central role in determining the performance of the user.

6 Modeling performance

In this chapter, a model is proposed that uses the data coming from the eye tracker to give an indication of the performance of the user during a certain task. It covers the part of the process titled "Performance of user is determined" in the flowchart depicted by figure 3. First, a pure mathematical model is given which assumes error free and ideal data, and uses this to construct a heuristic for performance. Then, the data is considered as it is in practice and modified such that it can fit into the defined model.

6.1 Mathematical model

To eventually be able to use the eye tracker for communication purposes, two aspects are important:

- Accuracy: how well the user is able to point their gaze in the right direction
- Fixation duration: for how long is the user able to keep their gaze in the correct location

This information has to be extracted from the data. First, the data should be defined and processed, to make sure the position of the gaze is obtained. Then, accuracy is defined and combined with fixation duration to construct a heuristic for the performance of the user.

First, the data is considered. This data is an sequence of samples, which we will denote by

$$(x_1, y_1),\ (x_2, y_2),\ ...,\ (x_N, y_N)$$

, where N is the number of samples collected, and x_i and y_i are the x and y coordinate of the i-th sample on the 2-dimensional coordinate system as defined in the previous chapter. The time between the samples is assumed to be constant and small enough such that the most important eye movements, the fixations, are captured in the samples.

6.1.1 Accuracy

The user is said to be correctly looking at an object in sample i if

$$(x_i, y_i) \in O$$

, where O is the target object as defined in 1. If this is the case, the accuracy of the gaze of the user is optimal, which we will define as 1. If the user is looking away from the screen, the accuracy is minimal, hence it will be 0 in this case. If the user is looking at some point on the screen outside the object, the accuracy is not optimal, so it should be lower than 1, but higher than 0, since looking at the screen is still better than looking away. One possibility for the measure of the accuracy in this case is to take the euclidean distance to the object minus the radius r of the object, and inverse it, to make the accuracy higher if the gaze of the user is closer to the object. Then we define the accuracy of the gaze of the user in a sample i to be

$$acc((x, y)) = \begin{cases} 1 & \text{if } (x, y) \in O \\ \frac{1}{1+|(x,y)-(x_o,y_o)|-r} & \text{if } (x, y) \notin O,\ 0 \geq x \geq w,\ 0 \geq y \geq l \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

Using this equation, the accuracy of a single sample can be determined. To get an idea of the overall accuracy of the user, the average accuracy of all samples can be calculated:

$$\frac{1}{N}\sum_{n=1}^{N} acc((x_n, y_n))) \tag{3}$$

6.1.2 Fixation Duration

For the fixation duration, a common term used is dwell time: the time gazed at an area of interest (AOI), in our case the object, from entry to exit. The total dwell time is the sum of all dwell times for a specific AOI over a trial [15]. Our AOI is the object, and the desired dwell time is the fixation time needed until the user is rewarded. Let this be denoted by t_a, the activation time.

To identify fixations, and to distinguish them from other events such as saccades, a possibility is to use an event detection algorithm, such as the one mentioned in chapter 2. The point of fixation could then be defined as the average of all processed samples in the fixation: this simplifies the model by averaging out the microsaccades occurring during the fixation. From this, the accuracy of this point of fixation can be calculated.

However, for this research, only the fixations are relevant. Also, the maximum sampling frequency of the used eye tracker is relatively low, namely 60 Hz. Therefore, such an algorithm is not appropriate for this research. Instead, the number of consequent samples which have an accuracy of 1, meaning the gaze was on the object, will be used to get an estimation of the fixation duration on the object.

If we denote the sampling frequency of the eye tracker by f_s, and consider the activation time t_a, then the amount of consequent samples that need to have perfect accuracy is

$$N_a = \lceil t_a * f_s \rceil \tag{4}$$

Fixations on the object are identified by the substrings, subsequences with only consequent samples, for which all samples in the substring have optimal accuracy and the substring length is maximal, meaning that the sample coming before and after the substring have suboptimal accuracy.

$$\begin{aligned} F = \{(x_i, y_i), ..., (x_{i+j}, y_{i+j}) \mid \; & acc((x_k, y_k)) = 1 \text{ for } i \geq k \geq i + j, \\ & acc((x_k, y_k)) \neq 1 \text{ for } k = i - 1, k = i + j + 1\} \end{aligned} \tag{5}$$

For every fixation f in F, let the length of the fixation, $\ell(f)$, be j, so the number of samples belonging to that fixation. A fixation has optimal duration if $\ell(f) \geq N_a$. An indication of performance with respect to fixation duration can be given by the fraction of fixations on the object that have optimal duration:

$$\frac{1}{|F|}\sum_{f \in F} \min(\left\lfloor \frac{\ell(f)}{N_a} \right\rfloor, 1) \tag{6}$$

6.1.3 Performance

To create a heuristic for total performance, a combination can be made of the performance with respect to accuracy and fixation duration. For example, a linear combination of the two equations 3 and 6 can be used, where the weight of both aspects can be changed as desired. However, also the product of the two could be considered. This would require both accuracy and fixation duration to be relatively large to ensure the overall performance is good, which is a quality that could be desired. Although both options can be helpful in some situations, the scores should also always be considered separately as well, as they can indicate what aspect the user struggles with or excels at.

6.2 Data in practice

The data does not always behave as expected in practice, as has been observed in chapter 4. Therefore, the data should be processed to fit into what has been described in the previous sections of this chapter.

One issue is that the samples are not completely accurate. There is possible bias in the data, as well as error or noise. Hence, the sample i can be written as

$$(x_i, y_i) = (X_i, Y_i) + (\beta_x(X_i), \beta_y(Y_i)) + (e_x, e_y) \tag{7}$$

where (X_i, Y_i) is the actual position of the gaze in sample i, $(\beta_x(x), \beta_y(y))$ is the bias in x and y, which was shown to depend on their own dimension, and (e_x, e_y) is the error or noise.

To estimate the position of the gaze in sample i, the error is neglected and the expression becomes

$$(X_t, Y_t) = (x_t, y_t) - (\beta_x^{-1}(x_t), \beta_y^{-1}(y_t)) \tag{8}$$

By applying this to all samples in the data, the bias will be removed to obtain a more accurate sequence.

The other problem is that in some samples the gaze could be undetected. In chapter 4, it was concluded that 5 consequent samples is a reasonable threshold for concluding the gaze was undetected during these samples. So, this holds for all samples part of a maximal substring of 5 or more samples with identical coordinates:

$$\begin{aligned} \{(x_i, y_i), ..., (x_{i+j}, y_{i+j}) \mid j \geq 5, (x_i, y_i) = (x_{i+1}, y_{i+1}) = ... = (x_{i+j}, y_{i+j}), \\ (x_{i-1}, y_{i-1}) \neq (x_i, y_i) \neq (x_{i+j+1}, y_{i+j+1})\} \end{aligned} \tag{9}$$

All samples contained in a sequence in the defined set need to be altered. In some cases, the decision could be made to remove all these samples from the data set. However, this could have an arguably unjustified positive impact on the accuracy. An alternative solution would be to set all these samples to the value $(w + 1, h + 1)$. The accuracy of the samples will then be 0, which can be justified, the fixation duration will not be affected, and no data is deleted. Hence, this is the suggested solution.

Finding the samples where the gaze is not caught by the eye tracker can happen after processing the samples by removing the bias, as this is a deterministic process, and hence identical samples in the original data will still remain identical after they have been processed.

7 Discussion

The main result of this research is the method to evaluate user performance in the give task. The most important parts are the heuristics for accuracy, fixation duration and for overall performance. All the proposed heuristics are based on literature, observations and real-life information provided by Koninklijke Visio. But unfortunately, no actual data was used to validate the proposed heuristics, due to several practical reasons. This however can still be done in the future, and could lead to interesting findings and implications. The lack of data makes it that the research is quite different from similar research, and difficult to compare, as it remains quite theoretical in that sense.

As seen in chapter 4, the data can exhibit unexpected behavior. Although this can depend on the eye tracker and software used, the data quality should always be investigated for eye tracker research, especially when the technology used is not the cream of the crop. Next to that, I would suggest always storing eye tracker data in the form of coordinates of the samples, possibly together with a time stamp but always in chronological order. With this, all relevant information can be extracted, for instance by using an event detection algorithm.

This research overall proposes a way to evaluate user performance in a simple eye gaze task. The actual value of the performance indicator may vary heavily among individuals, as individuals with PIMD are unique and their gaze behavior may show big differences. Therefore, each person needs to be assessed individually and on a case by case basis, under supervision of an professional. However, the evaluation can still offer value, especially when comparing performance of the same individual but for different occasions or moments. In this way, the findings of the research may still help professionals in assessing the progress in the ability of using an eye tracker of a person with PIMD.

8 Conclusion

In this paper, a method is proposed for evaluating performance in an eye gaze task for people with PIMD, based on the accuracy of their gaze and their ability to focus on the desired area for a given period of time. This can be used to help monitor the progress of individuals with PIMD during their training in using an eye tracking device. It can contribute to better training and can help maximize the amount of people with PIMD that can use an eye tracker in combination with a communication computer to enable them to better communicate with the outside world, including their caregiver. In this way, it will enable individuals with PIMD to express themselves and interact with the outside world more effectively.

9 Recommendations

In future research, the proposed evaluation of performance can be tested with an experiment, to check the validity of the model. Next to that, the model can be expanded or altered to enable the evaluation of performance of different tasks than the one chosen in this research, such that the whole training process can profit from the added value of user performance evaluation.

Also, the proposed performance indication can be used to apply the concept of Dynamic Difficulty Adjustment, as described in the background, to these training tasks: based on the performance evaluation, the difficulty of the task could adjusted in such a way that the task is neither too easy nor too difficult. A possibility could be to determine user expertise by the quantified performance and by a quantification of the task difficulty, and to perform regression on the expertise to estimate the user expertise in the future and adjusting the task difficulty to achieve optimal performance. With this, the training could be useful for a larger number of individuals and could become more effective.

References

[1] J. Mansell. Raising our sights: services for adults with profound intellectual and multiple disabilities, Mar 2010.

[2] H. Nakken and C. Vlaskamp. A need for a taxonomy for profound intellectual and multiple disabilities. *Journal of Policy and Practice in Intellectual Disabilities*, 4(2):83–87, 2007.

[3] M. Cigale, M. Lustrek, M. Gams, T. Krämer, M. Engelhardt, and P. Zentel. The quest for understanding: Helping people with pimd to communicate with their caregivers. 10 2018.

[4] R. Andersson, M. Nyström, and K. Holmqvist. Sampling frequency and eye-tracking measures: How speed affects durations, latencies, and more. *Journal of Eye Movement Research*, 3:1–12, 09 2010.

[5] K.M. Wilkinson and T. Mitchell. Eye tracking research to answer questions about augmentative and alternative communication assessment and intervention. *AAC: Augmentative and Alternative Communication*, 30(2):106–119, 2014.

[6] G.S. Townend, P.B. Marschik, E. Smeets, R. van de Berg, M. van den Berg, and L.M.G. Curfs. Eye gaze technology as a form of augmentative and alternative communication for individuals with rett syndrome: Experiences of families in the netherlands. *J Dev Phys Disabil*, 28:101–112, 2016.

[7] K. Vessoyan, G. Steckle, B. Easton, M. Nichols, V. Mok Siu, and J. McDougall. Using eye-tracking technology for communication in rett syndrome: perceptions of impact. *AAC: Augmentative and Alternative Communication*, 34(3):230–241, 2018. cited By 14.

[8] K. Topper and W. Huls. Projectaanvraag sprekende ogen 2.0, 2020.

[9] R. Garris, R. Ahlers, and J.E. Driskell. Games, motivation, and learning: A research and practice model. *Simulation and Gaming*, 4(4):441–467, Dec 2002.

[10] K. Rayner. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3):372–422, 1998. cited By 4403.

[11] M. Nyström and K. Holmqvist. An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. *Behavior Research Methods*, 42:188–204, 02 2010.

[12] M. Zohaib. Dynamic difficulty adjustment (dda) in computer games: A review. *Advances in Human-Computer Interaction*, 2018, 2018. cited By 25.

[13] Lawrence A. Beck. Csikszentmihalyi, mihaly. (1990). flow: The psychology of optimal experience. *Journal of Leisure Research*, 24(1):93–94, 1992.

[14] R. Hunicke and V. Chapman. Ai for dynamic difficulty adjustment in games. volume WS-04-04, pages 91–96, 2004. cited By 136.

[15] K. Holmqvist, M. Nyström, and F. Mulvey. Eye tracker data quality: What it is and how to measure it. pages 45–52, 2012.