HOW PEOPLE WITH A VISUAL FIELD DEFECT SCAN THEIR ENVIRONMENT: AN EYE-TRACKING STUDY

Masterthesis Applied Cognitive Psychology
27.5 ECTS

Henrike Haase
6545424
H.Haase@students.uu.nl

Utrecht University
First Supervisor
Krista Overvliet
K.E.Overvliet@uu.nl

Second Supervisor
Nico Romeijn
N.Romeijn@uu.nl

Royal Dutch Visio
Jan Koopman
JanKoopman@Visio.org

Abstract

The scanning behavior of people with a Visual Field Defect (VFD) is still relatively unexplored, although a VFD can have great impact on the person concerned with it. This study examined the scanning behavior of people with a VFD in a mobility situation by the means of an Eye-Tracker.

Participants were asked to watch a video of everyday mobility situations, showing both walking and biking sequences from first person perspective. They were asked to imagine being that person in the video and look around naturally. In the video, 33 objects were defined as regions of interest (ROIs), as they are important to detect in order to walk or bike safely. The Eye-Tracker recorded when and where the participants looked. It was expected, that people with a visual field defect (VIPs) need more time to fixate on relevant ROIs for the first time, detect less ROIs in total, do not look at them as long as sighted people do and fixate more often on the ROIs.

The results confirmed that VIPs detect less ROIs and spend less time looking at them. Contrary to what was expected, the results also showed that sighted people fixate more on the ROIs than VIPs. An additional analysis showed that the time until an ROI is first looked at, is significantly shorter for sighted people.

These findings showed that there are indeed differences in scanning behavior and they could be used to help VIPs counterbalance their decreased visual field.

When moving around, people look at many things, like billboards, shop windows or at other people. Meanwhile, without much thinking about it, they also scan their environment for potential obstacles like pillars, the edge of a sidewalk or other road users. Scanning the surrounding and thereby detecting potential obstacles happens rather intuitively (Pambakian, Wooding, Patel, Morland, Kennard, & Mannan, 2000). If something in the visual field might hinder moving around safely, it is perceived and the behavior is adapted accordingly. This does not only happen for objects that are in the center, but also in the periphery. But what happens if people have a visual impairment that restricts the visual field? A visual field defect (VFD) is an impairment in which people miss parts of their visual field, resulting in an incomplete overview (Pollock et al., 2011). In this case, road users or other barriers might pose a risk if they are not seen in time.

Research has shown, that visually impaired people (VIPs) have indeed significantly more problems to see relevant obstacles and people in time (De Haan, Melis-Dankers, Brouwer, Tucha, & Heutink, 2016). One possible explanation is a less efficient way of scanning the environment for important objects (Pambakian, Wooding, Patel, Morland, Kennard, & Mannan, 2000). This could lead to a delay to detect potential threads or a difference in how long an important object is looked at. Although the consequences of not seeing a threat in time can be severe, there has not been a lot of research on scanning behavior of people with a VFD in mobility situation. Thus, it is worth investigating if there is a difference in scanning behavior between people with and without a visual field defect in a mobility situation.

The human visual field corresponds to 180° horizontally and 130° vertically (Duchowski, 2007). There are different regions within the visual field, each fulfilling a different task. Vision is sharpest around 2° circular region around the center of the gaze and quickly declines after that, with a sharp drop in acuity after 5° (Duchowski, 2007). The so-called useful field of view is the area, in which information can be perceived at a glance without having to move either the head or the eyes (Edwards, Ross, Wadley, Clay, Crowe, Roenker, & Ball, 2006). This area extends to around 30°. To see objects further in the periphery in detail, eye- or head movements have to be made. Although the periphery is less sharp, it is important for motion detection and serves as a warning system for moving objects in the visual field (Duchowski, 2007). This is important when objects have to be detected quickly, for example when riding a bike or driving a car. For driving a car, the European requirement is a visual field of at least 120° (Papageorgiou, et. al., 2007).

A VFD can manifest itself in different forms. For example, people suffering from a complete hemianopia miss one half of their visual field. Others suffer from quadrantanopia, where they miss a quarter of the visual field (de Haan, Heutink, Melis-Dankers, Tucha, & Brouwer, 2014). Other forms of visual field loss include peripheral vision loss, e.g. retinitis pigmentosa, resulting in a "tunnel vision" or central field loss, e.g. macular degeneration, leading to impairments in the center of the visual field (Petzold, & Plant, 2005; Warren, 2009).

Research has shown, that a VFD can lead to difficulties with daily activities, like using a phone, preparing a meal and can have a huge impact on quality of life in general (Williams, Brody, Thomas, Kaplan, & Brown, 1998; de Haan et al., 2015). Furthermore, it was shown, that a VFD is related to reduced mobility performance (Pollock et al., 2011; Black, Lovie-Kitchin, Woods, Arnold, Byrnes & Murrish, 1997) and an increased number of incidents, like tripping or bumping into obstacles (Coeckelbergh, 2002). Common mobility problems include detecting relevant obstacles in the (restricted) periphery and seeing people or obstacles in time (De Haan, Melis-Dankers, Brouwer, Tucha, & Heutink, 2016; de Haan, et al., 2015), especially when the surrounding is unfamiliar (de Haan, 2014, Pollock et al., 2011). Interestingly, problems with regards to mobility have only been shown for people with a VFD, but not for people with acuity problems, which is the "ability to resolve detail" (Colenbrander, 2002; Black, et al., 1997). For example, research on elderly drivers has shown that impairments in the visual field are related to higher collision rates, whereas acuity problems are not (Kwon, Huisingh, Rhodes, McGwin Jr, Wood, & Owsley, 2016; Coeckelbergh, 2002). Acuity is needed for tasks involving details like reading (Colenbrander, 2002) and can be more easily compensated, for example with glasses and lenses.

So far, only the consequences of a decreased visual field have been investigated, but not the underlying viewing pattern of people with a VFD. As mentioned, their scanning behavior might be less sufficient, as suggested by research using static pictures for visual search tasks (Pambakian, Wooding, Patel, Morland, Kennard, & Mannan, 2000).

When people scan their environment, they quickly direct and fixate their gaze at what attracts their attention (Eckstein, Guerra-Carillo, & Singley, 2017). According to literature, a fixation is made when the same gaze position is kept for 100ms (Lee, Black, Lacherez, & Wood, 2016). Fixations are indicative of when available information was perceived (Olsen, 2012). To investigate this, the time until fixation, their duration and location have been used in previous research (Nilsson, & Nivre, 2011; Coeckelbergh, 2002). In between these fixations, they make rapid eye movements, called saccades. These saccades can be both controlled (towards a target) and uncontrolled (e.g. stimulus-driven) (Duchowski, 2007). During these saccades, none or only little visual information is processed (Salvucci, & Goldberg, 2000).

Insufficient scanning behavior is displayed as “uncontrolled and time-consuming gaze-shifting" (Zihl, 1995; Warren, 2009) and an increase in total amount of fixation (Warren, 2009; Papageorgiou, et. al., 2007) and saccades (Lee, Black, & Wood, 2017). This is similar to the scanning behavior of people without any visual impairment in complex situations, where people make more fixations and show longer scanpaths (Eckstein, Guerra-Carillo, & Singley, 2017).

Additionally, research on people with hemianopia has shown, that gaze is shifted to the intact side first and with significantly higher number of fixations both in the intact and affected hemifield. It seems that they have difficulties to organize their search strategy, especially in the affected area (Zihl, 1995). Consequently, less time is spent on the relevant objects, which makes it harder to comprehend the visual information and act appropriately (Warren, 2009). So far, this has only been shown in static tasks (Pambakian, Wooding, Patel, Morland, Kennard, & Mannan, 2000). Little is known about how people with a VFD scan their environment when they move around.

It is important to investigate how mobility problems arise, as they pose great safety risks for the people involved. When walking or driving, objects have to be detected quickly and information comes from many different sources. If parts of the visual field are missed, people or objects can appear unexpectedly in the visual field. Furthermore, it is necessary to understand what goes wrong in these situations to develop more beneficial training sessions.

The aim of this study is to investigate the differences in efficiency of scanning behavior between sighted people and VIPs in a mobility situation.

The added value of this study is not only the investigation of scanning patterns of people with a VFD in a mobility situation, but also the use of a wearable Eye-Tracker.
Eye-Tracking has already been used in research concerning reading, problem solving or advertisements (Salvucci, & Goldberg, 2000). This is often done with remote Eye-Trackers, which are fixed to a computer screen or use a chin rest, both of which restrict movement. Newer Eye-Trackers are wearable and head-mounted, allowing more natural behavior and record the gaze relative to the subject (MacInnes, Iqbal, Pearson, & Johnson, 2018). Eye-Tracking might offer great insight into scanning behavior of people with a VFD as it allows to measure more precisely where people look and for how long. If scanning behavior is indeed different between people with and without a VFD, it could be made visible with the help of the Eye-Tracker.

Based on the studies mentioned above, it is expected that there are indeed differences in scanning behavior which can be investigated in more detail by means of an Eye-Tracker. These impacts are expected to influence both the time and amount of detections and fixations. As described, reduced acuity does not influence scanning behavior in a mobility situation, which is why no distinction was made between VIPs with and without acuity problems. The focus of this study is on the difference in scanning behavior of people with a field defect, irrespective of acuity deficits.

Earlier research showed, that VIPs have trouble to detect an obstacle and fixate later on them (Lee, Black, & Wood, 2017). Thus, it was expected that VIPs take more time to detect an obstacle than people without a VFD. For VIPs, scanning behavior is less sufficient, resulting in an overall lower rate of object detection (Pambakian, et al., 2000). It was therefore expected that VIPs look at less ROIs. Furthermore, more saccades are made to capture the whole scenery, leaving less time to spend on each object (Lee, Black, & Wood, 2017). Thus, dwell time on each ROI was expected to be reduced for VIPs. As stated above, generally more fixations have to be made to comprehend the visual image (Warren, 2009). This also leads to a higher number of total fixations on the ROIs made by VIPs.

Method

Participants

The study was conducted at Royal Dutch Visio, which is a center of expertise for the blind and partially sighted. In total, 32 participants were tested (14 female, 18 male; mean age = 42.46, SD = 20.05). Twelve participants were fully sighted (7 female, 5 male; mean age = 32.67, SD = 15.48). They were recruited from Visio and the University of Utrecht and had a self-reported uncorrected visual acuity of at least 0.5 and/or corrected visual acuity of 1.0. The sighted participants were used as a reference. Twenty participants had a visual field defect (7 female, 13 male; mean age = 52.26, SD = 19.14). Of these 20 VIPs, one had to be excluded, because the data was not saved properly and two others had to be excluded due to too much missing data. Participants with a visual impairment had an appointment with an employee from Visio in Amsterdam, e.g. with an occupational therapist or ICT trainer, for a counselling or training session.

Although acuity deficits are supposedly not of any influence, it was checked that they were still able to identify objects in the surroundings. For the analysis, VIPs were not only considered as one group, but their type of field defect was also taken into consideration. Namely, a distinction was made between a central field defect and a peripheral field defect. As research has shown, the periphery is of importance to detect moving objects at an early stage and has been shown to pose other problems with regards to mobility than a central defect (Coeckelbergh, 2002). The VIPs might therefore differ in their scanning behavior.

Unfortunately, the medical conditions of the VIPs were not always straightforward. Hemianopia for example can be considered as both a central and a peripheral field defect. Also, the sizes of their central and peripheral field loss differed to a great extent. The allocation of VIPs to one of the two groups was made with the help of a professional from Visio.

Materials

Eye-Tracker

The Eye-Tracker was a binocular Eye-Tracker by Pupil Labs. For each participant, the Eye-Tracker had to be calibrated. Research has shown, that a minimum of twelve fixation points already offers good calibration quality (Chen, Tong, Gray, & Ji, 2008). As a good calibration is crucial for high quality data, it was chosen to add three more fixation points and evenly distribute them over the screen (MacInnes, Iqbal, Pearson, & Johnson, 2018). Thus, numbers from one to 15 were displayed at the large screen, ranging from upper left to lower right. Pupil Labs uses three cameras, one directed at each eye and another one directed at the environment, in the direction of the participants' gaze. The cameras directed at the eyes ran on 200Hz, whereas the world cameras ran on 30Hz. The combination of an outward facing world camera, that captures the subject's point of view, and two cameras directed at each eye respectively, which record the pupils' position allow a more natural recording, because the participant can move freely (Duchowski, 2007). Information from both sources is considered when the subjects' gaze is recorded and displayed.

Stimulus Video

The stimulus video had been recorded previously by Visio at 30 frames per second (fps) and lasted around six minutes. It was shown on a large screen (1,08m x 1,77m), where luminance was kept around 130-140 cd/m^2, in order to control for effects of difference in lighting levels, which have been shown to influence performance on mobility tests (Black, Lovie-Kitchin, Woods, Arnold, Byrnes & Murrish, 1997). The stimulus-video showed everyday situations, with short sequences on a bike and sequences of walking in a shopping centrum. Obstacles that had to be detected during the bike sequence included the edge of the pavement, cars and other cyclists approaching from the sides as well as pedestrians. The shopping center additionally posed other obstacles like street lanterns. Both sequences included moving and static obstacles.

Obstacles in the video were chosen based on a poll questionnaire with 30 professional mobility instructors who rated an extensive list of potential obstacles shown in the video. They had to be rated on a Likert scale of one to ten in terms of how important they were to detect in order to move around safely. From this, all objects with an average score above eight were chosen, which resulted in 33 objects.

Procedure

Participants were first informed about the purpose of the study, that all information was collected anonymously and that their participation could be stopped at any given time. Afterwards, they signed the informed consent forms. Participants were then seated in front of a large screen at one meter distance and put on the Eye-Tracker so that it was comfortable for them and they could move their head freely. If they did not have any previous experience with Eye-Trackers before, the method was explained to them. In Pupil Capture, the correct position for each eye had to be adjusted, which was done according to the instruction provided by Pupil Labs. Afterwards, the Eye-Tracker was calibrated by fixating on the 15 fixation points on the screen. The participants were instructed to fixate each number while the experimenter clicked on the corresponding number on the laptop screen, in order to “match” the gaze to the position on the big screen. During this, participants were not allowed to move their head. While watching the video, participants were instructed to look naturally at the screen, move their head if needed and pretend to be the person walking and biking in the video.

Figure 1. Setting of the study. The boy on the bike and the pedestrian are objects to be detected

Design

The condition of vision was the independent variable. People were labelled as either sighted or VIP, when two groups were compared. In the case that VIPs were divided into two different conditions of a VFD, they were either assigned to central or peripheral field defect.

The dependent variables are the time until first fixation, number of detected objects, dwell time and number of fixations on the objects in the video.

Analysis

Software

For the Eye-Tracker, software from Pupil Labs was used, namely Pupil Capture and Pupil Player. Pupil Capture was used for the calibration of the Eye-Tracker, to set the parameters regarding the recording and to record the gaze. Pupil Player was used to playback the video and to export the recordings and gaze data. Data preparation, noise reduction and gaze analysis were done using PyCharm which works on Python 3.5, while statistical analysis was done with SPSS 25.

Data preparation

For data extraction, first some preparations had to be made to the video and in the Pupil software. In total, ten QR-codes were placed around the borders of the video which were displayed once it started and lasted until the end of it (see figure 1).

These QR-codes served two purposes. First, they defined the surface on which the video was displayed. While these QR-codes are visible, Pupil Capture groups the data together in one file. Also, it translates this surface into a coordinate system ranging from (0,0) at the lower left corner to (1,1) at the upper right. The corners of the surfaces' coordinate system were determined by the black edges of the QR-codes, therefore cutting a small frame around the video, that was not considered (see figure 1). This was later corrected for. Every gaze position was then calculated based on this coordinate system. Thus, there is no reference to any real-world object, only positions in the coordinate system. Second, the QR-codes marked the beginning and the end of the stimulus video in the recording. Therefore, it was possible to calculate when each obstacle in the video was visible, relative to the starting point of the stimulus video.

As described above, the 33 most important obstacles were defined by experts of the field. These “regions of interest” (ROIs) were manually defined for each object in the original stimulus video. It would have been too time-consuming to determine the ROIs on each frame in the video with a length of almost six minutes and 30 frames per second (fps). Instead, it was chosen to take a picture of the movie every second and determine the coordinates of the ROIs in there. The intermediate ROIs were calculated based on the frame rate, meaning that the positions in between two seconds were calculated by the coordinate change from one picture to the next. These pictures were also used to determine their presentation time in the stimulus video. To check if the ROIs matched the intended objects, all ROIs were printed on the pictures in the video, which resulted in roughly 10.200 pictures of which 4.500 had one or more ROIs. These pictures were controlled and if necessary, adjustments were made to the coordinate info for each ROI.

Finally, some adjustments had to be made to the coordinate information, because it used a different system than the gaze data did. The pictures relied on pixels, which is why the coordinates of the ROIs ranged from (0,0) in the upper left to (1920,1080) in the lower right corner. This coordinate system also took the frame around the video into account, which was neglected by the Pupil software. Therefore, this frame (with a width of 18 pixels on each side) had to be subtracted from each side. Afterwards, all x and y pixel positions were normalized by dividing them by 1884 (1920 – 2x18) or 1044 (1080 – 2x18) respectively. The last step was to mirror along the y-axis, in order to align the pixel coordinate system with the gaze coordinate system from lower left to upper right. These steps were necessary in order to translate the gaze position information onto the coordinates of the ROIs.

Noise reduction

In Pupil Player, a file could be exported for each participant with information on every recorded gaze. This included timestamps, the gaze positions in the coordinate system, the "world frame" number and the level of confidence of that gaze position. The world frame indicated gaze positions in terms of their referring frame from the world camera video which ran on 30Hz. The level of confidence ranged from 0.0 (low confidence) to 1.0 (high confidence). This value served as a criterion for data quality as it was determined by the quality of the pupil detection. The better the detection of the pupil worked, the higher the level of confidence was. A good calibration was necessary for sufficient pupil detection and in turn, high confidence data. To ensure data quality, participants were asked to again fixate on some points after calibration in order to check for correspondence between the fixation and the marker. In case of poor correspondence, which was determined by the researcher, the calibration was repeated until correspondence was sufficient.

The first step to reduce noise concerned the number of gaze positions. Depending on lighting and CPU, it differed how many gaze positions per second and world frame were calculated by Pupil Capture. On average, there were about 250 entries per second. Due to unknown reasons, Pupil Player occasionally produced data sets with thousands of entries for one world frame. To control for this error, a maximum of 20 gaze data entries per world frame was set. Entries per world frame differed, but were usually below 20. Choosing that as the maximum therefore offered a good trade-off between keeping calculated gaze-positions and excluding flawed data.

As mentioned, the accuracy of the gaze data differed in terms of their confidence levels, ranging from 0.0 to 1.0. Also, the amount of missing data differed not only between sighted people and VIPs, but also within the VIP group, which made it difficult to find common grounds and determine reasonable cut-off points and thresholds. Sometimes, parts of data sets had very low confidence, although their gaze points logically followed each other. In other cases, however, there was higher confidence for unreasonable data, e.g. values that were off the surface. Therefore, interpolation at a certain confidence level was necessary. Cut-off scores were either too strict and resulted in high interpolation rates, whereas lower confidence levels took too much unreasonable data into account. It has to be stressed, that interpolating the data influences the analysis to a great extent, as it might prolong or shorten a sequence that a participant looks at a ROI. Therefore, a confidence level of 0.6 was set as it offered the best trade-off between retaining original data while at the same time filling missing data points. Gaze positions below that confidence level were interpolated by the two adjacent valid data points. This way, missing data points, for example due to blinking, could be filled easily, making data more robust against these disturbances. However, if the interpolation period exceeded 75ms, data was not considered as it was not reliable enough. This “maximum gap length” has been practice in other Eye-Tracking studies as well (Olsen, 2012).

Because of these adjustments, two of the VIPs had to be excluded from the analysis. Eighty to 90% of their gaze data was below the confidence threshold of 0.6, resulting in interpolation periods of more than three minutes (188.34s and 202.38s, respectively). Other research also excluded participants, where more than 50% of the data was missing (Lee, Black, & Wood, 2017).

Matching process between gaze and ROI

To conclude whether an object was looked at and for how long, timestamps from both the ROI and the gaze data were compared and the gaze positions from the Eye-Tracker were translated onto the coordinates of the video frames. In case the gaze fell in the ROI, it was counted as a “match”. From this, the time until the first match, total length and number of all matches (in seconds) were extracted. If the participant’s gaze and the ROI matched for longer than 100ms in a row, it was counted as a fixation (Lee, Black, & Wood, 2017). The time until first fixation and the total number of fixations were taken. Additionally, it was calculated how many ROIs of the 33 in the video were seen.

Results

In this study, differences in scanning behavior between sighted people and VIPs were investigated. VIPs were investigated both as one group and also separately, depending on their type of VFD, which was either central or peripheral. An alpha level of .05 was used for all statistical analysis.

Descriptive Statistics

Sighted people detected 26.25 ROIs on average (SD = 2.41), compared to 22.23 (SD = 4.98) detected by the VIPs, specifically 21.64 (SD =1.14) by VIPs with a central defect and 23.34 (SD = 4.76) by VIPs with a peripheral defect. Minimally, participants detected 12 and maximally 30 ROIs. 23.34 (SD = 4.76)

Calibration for the people with a visual field defect was often more difficult and had to be done more than once in order to get a sufficient result. This influenced the amount of data with a low confidence level. The average interpolation period for VIPs was higher than the one for sighted people, the same is true for the average interpolation period above the threshold of 75ms and the average percentage of data below the confidence threshold of 0.6.

Table 1

Descriptive Statistics. Standard deviations are given in between brackets

Participants	N	Average interpolation period in seconds	Average interpolation period above threshold in seconds	Data below confidence level in percent
Sighted	12	26.41 (32.99)	9.68 (7.09)	10.07 (9.72)
VIP	17	28.21 (25.36)	17.23 (15.49)	13.84 (11.02)
Central defect	11	25.08 (20.09)	16.74(14.53)	12.07 (9.04)
Peripheral defect	6	33.95 (34.49)	18.74 (18.50)	15.94 (14.72)

Time Difference in Object Detection

To determine when objects were detected on average by sighted people and VIPs, a survival analysis was conducted. Survival analysis is used to determine the difference until an event happens, meaning the time required to detect a ROI. To determine this difference, the mean time until the first fixation was used as the dependent variable. Thus, the time until the participant looked at the ROI for at least 100ms. It was expected that VIPs take more time to detect a ROI. A log rank test was conducted to determine differences between VIPs and sighted people. The distributions for the two groups was not statistically different $\chi^2(1) = 0.415$, *p* = .519. When the VIPs were split into groups regarding the location of their field defect (peripheral and central), the result was also not significant $\chi^2(2) = 3.184$, *p* = .203.

Total Object Detection

A Poisson regression was used to investigate the difference of how many objects were detected in total by sighted people and VIPs. It was expected, that VIPs detect less objects due to their less controlled scanning strategy. A detection was given, when there was at least one match for an ROI. Sighted people detected 18% more objects than VIPs (95% CI, 1.017 to 1.371), which is a significant result *p* = .030.

However, when the analysis was run with two groups of VIPs, the result was not significant *p* = .074.

Dwell Time

It was hypothesized, that the dwell times differ between sighted people and VIPs, namely that VIPs spend less time looking at the ROI. The dependent variable dwell time was expressed as the mean of the total length of all matches. To test this hypothesis, a normality check was done first. Although the Shapiro-Wilk test indicated a normal distribution (*p* = .769), Levene's test of homogeneity of variance showed that the distribution was not normal (*p* = .738). Therefore, non-parametric tests were chosen.

To compare sighted people and VIPs, the Mann-Whitney U test was used. It can be concluded that VIPs spend less time looking at the ROIs (U = 50.000, *p* = .021).

To compare dwell time between sighted people and the different types of VFD, the Kruskal-Wallis test was used, which showed that there was no significant difference in dwell time between the three groups $\chi^2(2) = 5.550$, $p = .062$.

Overall Number of Fixations

It was expected that the overall number of fixations differ between sighted people and VIPs, with VIPs having a higher number of fixations. First, a normality check was done. The Shapiro-Wilk test indicated a normal distribution ($p = .612$), but Levene's test of homogeneity of variance depicted a non-normal distribution ($p = .181$), which is why a non-parametric test was chosen for this hypothesis. The Mann-Whitney U test was used to compare sighted people and VIPs. It showed that sighted people exhibited a higher number of fixations ($U=38.5$, $p =.004$). Therefore, the hypothesis was not confirmed.

To compare sighted people and the two types of VFD, a Kruskal-Wallis test was used. It showed a significant difference in number of fixations between at least two groups $\chi^2(2) = 7.962$, $p = .019$. A post-hoc test using Bonferroni correction was administered to test pairwise comparison (see figure 2). It showed that there was a significant difference between sighted people and VIPs with a central defect $p = .043$.

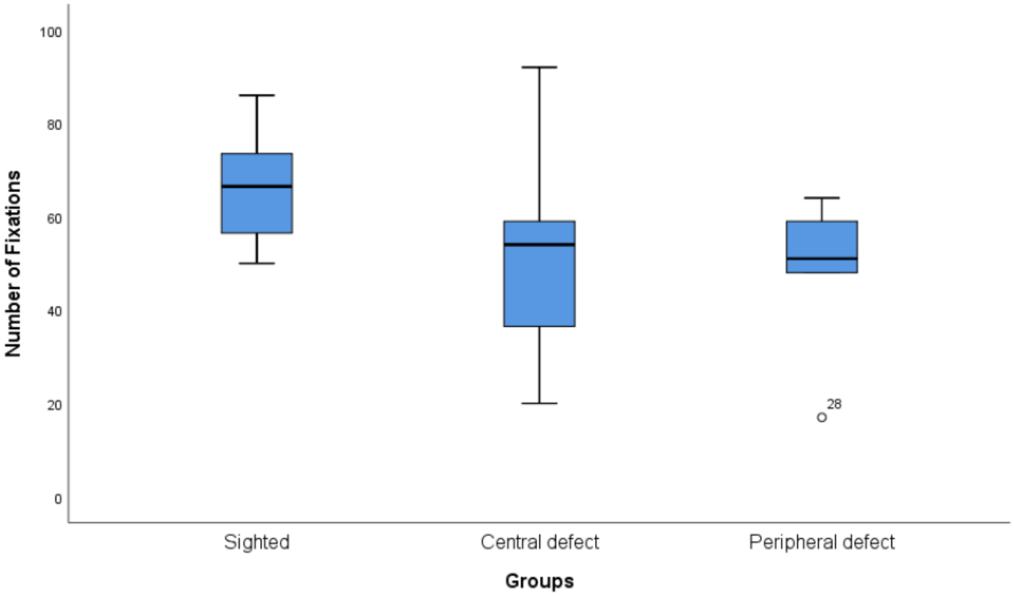

Figure 2. The number of fixations for each group of participants. Only the difference between sighted people and VIPs with a central defect was significant ($p = .043$)

Further investigation of Time Difference in Object Detection

Usually, fixation measurements are used to show that somebody actually saw a relevant object, as we assumed in the first hypothesis. However, this could pose a problem, as it is still debatable if somebody really perceived an object after 100ms (Duchowski, 2007). It could also be the case, that more time is needed to understand complex situations. For easier tasks, such as detecting a (moving) object, the opposite could be the case. Therefore, an additional analysis was conducted, using the time until the first match between gaze and ROI as the dependent variable. This result was significant $\chi^2(1) = 5.004$, *p* = .025. The gazes from sighted people "matched" the ROIs earlier than the gazes from the VIPs (M = 1.560s, SD = 0.154 and M = 2.110s, SD = 0.379; see figure 3). The result was not significant, when VIPs were split into two groups (*p* = .068).

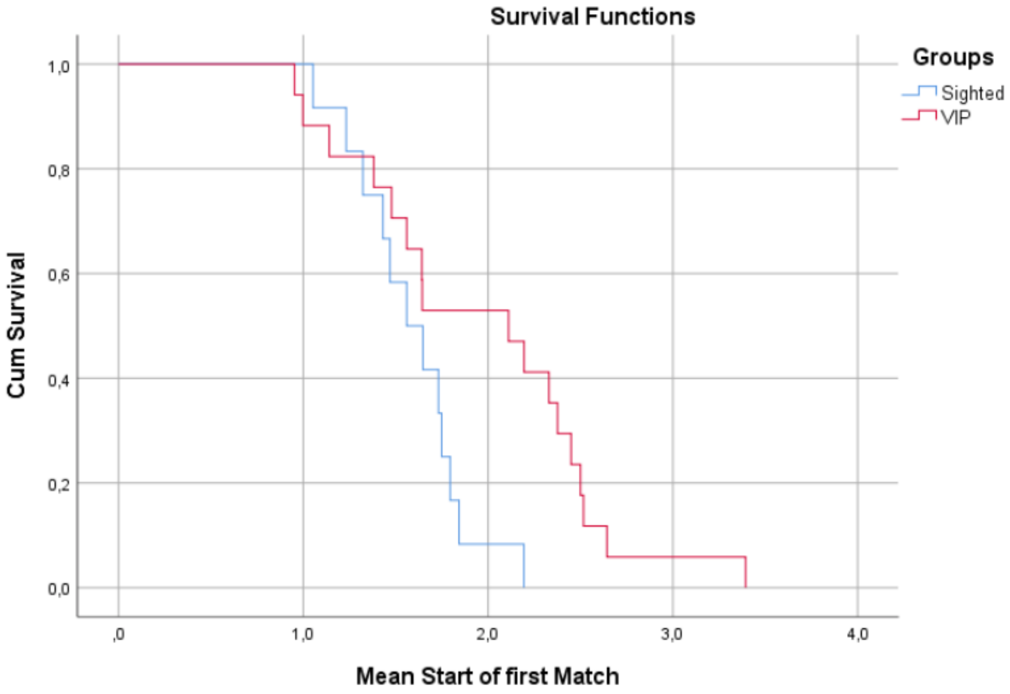

Figure 3. Survival functions for the time until the first match was made on a ROI. Sighted people matched earlier than VIPs.

Discussion

This study showed that VIPs detect less ROIs and spend less time looking at them. The results also showed that sighted people fixate more on the ROIs than VIPs, which was not in line with the hypothesis. Additionally, it was shown that the time until an ROI is first looked at, is significantly shorter for sighted people. This means that there are indeed differences in the efficiency of scanning behavior between sighted people and VIPs.

It was hypothesized that VIPs detect potential obstacles later than sighted people as research has already shown (de Haan, et al., 2015). However, the results did not confirm this. In this study, the time until the first fixation was used to determine whether participants detected a ROI. The definition of a fixation was crucial for this. Here, fixations were defined as consecutive gaze-data points on the ROI. It could also be that people often scan around the “edges” of a ROI, which could be sufficient for detection. However, if one gaze point fell outside the ROI, it was not counted as a fixation anymore, but rather as a new match.

Further, it was expected that sighted people detect more objects in total (de Haan, et al., 2015). This was also the case in this study, even though the sighted people also did not look at every ROI. Possibly, some objects like the edge of a pavement were not looked at directly, but were only perceived via the periphery. For sighted people, this ROI might usually not pose a risk and is therefore not looked at specifically. The selection of objects was mainly based on safety relevance. Therefore, this result stresses even more how VIPs are more at risk of missing important obstacles.

As expected, VIPs spend less time looking at the ROIs. The literature has shown, that this could be due to their unorganized search strategy (Zihl, 1995; Warren, 2009). This could also be the root of the higher rates of incidents for VIPs in mobility situations. Regardless of whether a fixation (with a certain length) is made at an obstacle, the shorter dwell time could indicate that VIPs are too busy scanning their environment instead of being able to focus on single relevant obstacles. Research has shown, that VIPs often try to overcompensate for their VFD by scanning more to their affected field in general (Lee, Black, & Wood, 2017). Unfortunately, this has the consequence of spending less time on the more relevant objects. As less information about the objects is available, they might be more surprised by their behavior, which could cause a feeling of uncertainty.

It was expected, that the number of fixations is higher for VIPs (Pambakian, 2000; Warren 2009). However, the results did not confirm this. On the contrary, it showed that sighted people make significantly more fixations than VIPs, especially compared to VIPs with a central field defect. In line with significantly shorter dwell time for VIPs, their unorganized search strategy could be a possible explanation for this. A fixation is made only after 100ms, which might be difficult to achieve if scanning is more unorganized and involves more saccadic movements. They do not spend as much time on a ROI, let alone fixate on it. The location of the VFD could also play an important role here. In order to see an object that is in the center of the visual field, fixations have to made next to the actual ROI and not directly at them. Therefore, important information about an object might be missed, resulting in more incidents.

There are limitations of the study that need to be considered. First of all, the screen size and the video itself already determined the visual field of the participant. Participants said for example, that they wanted to “look” into the streets to their right/left at an intersection, but were not able to as it was at the end of the screen or because the recording did not show it. In a real driving situation, they might have looked differently and therefore detecting other objects. On the one hand, by turning their head, they might have detected objects from the side and at the same time missed information on what was in front of them. On the other hand, turning the head might have also posed a chance for VIPs with a central defect to detect objects in front, as they were not occluded by their field defect anymore. Also, no distinction was made between the location of the objects in the video. VIPs with a peripheral defect might have more difficulties with objects approaching from the sides. VIPs with a central defect might have more difficulties with objects that they encounter in front of them. However, this was difficult to determine in this video, as the ROIs often moved through the whole screen and were not easily classifiable as either a central or peripheral ROI.

Moreover, there were some difficulties regarding the use of the Eye-Tracker. Pupil Labs relies on the difference between the pupil and the iris for detection. It was challenging when participants were wearing glasses, when they wore lenses or had a disease that influenced pupil color, like a cataract. Data for VIPs had to be interpolated more often and for slightly longer periods than the data from sighted people.

There were also strong points to this study. The video showed scenes that were recorded in the streets and in a real shopping center. The ROIs indicated objects of different shapes and sizes, for varying amounts of time, were both static and moving and came from different sides. Participants said that they experience similar situations like the ones in the video, e.g. cyclists appearing unexpectedly or obstacles like pillars and advertisements in their way. Therefore, the stimulus video was realistic and conclusions can be taken into the real environment of VIPs.

To our best knowledge, this study was one of the first to investigate the scanning behavior of people with a visual field defect in a mobility situation regarding walking and biking. So far, there have been studies on scanning behavior, often either related to specific field defects or to other situations like driving. Similar to previous studies (e.g. Pambakian et al., 2000) the scanning behavior of fully sighted people was used as a reference. Thus, it is ensured, that VIPs are compared to a realistic standard, as fully sighted people also do not detect every obstacle, e.g. due to too much cognitive load (Eckstein, Guerra-Carillo, & Singley, 2017).

Furthermore, the use of a head-mounted Eye-Tracker was also a valuable addition to study. With the use of a mobile Eye-Tracker, compared to a remote one, a bigger visual field can be assed in a realistic setting and scanning behavior can be measured from the person's point of view.

In summary, it was shown that there are differences in the speed of detecting an object, how long and how many objects are looked at. The problems that VIPs encounter in mobility situations could result from these differences in scanning behavior, which means that institutions like Visio should focus on counterbalancing these shortcomings in order to help VIPs to move around more safely.

Application and recommendation for Visio

The results of this study and the use of an Eye-Tracker have great potential in facilities like Visio. In future training sessions, the insight into the scanning behavior of a patient could be used by the therapists to get a better understanding what potential problems are. Also, patients could be able to get more self-aware when they look back at the video and see their own scanning behavior displayed. Spouses and other family members can also benefit as they get a clearer understanding of how that person perceives a scenery.

Furthermore, the Eye-Tracker could be implemented as a sort of control mechanism for insufficient scanning behavior. Previous research has already shown, that there is a link between what is relevant for a task and what is looked at (Hayhoe, & Ballard, 2005). In a mobility situation, this could be used to control for and enforce a more efficient scanning strategy.

In the future, different levels of stimuli difficult could also be used to determine how behavior changes in highly cluttered situations. Studies on scanning behavior could also go one step further in Eye-Tracking with the use of Virtual Reality (VR). The fact, that the screen forces the participant to look in front already poses a restriction in their natural scanning behavior, especially because VIPs tend to look down while walking, according to experts at Visio.
Three of the participants also use a cane in their daily life for cluttered mobility situations, which of course they were not able to use during the experiment, but could be in VR.

References

Black, A., Lovie-Kitchin, J. E., Woods, R. L., Arnold, N., Byrnes, J., & Murrish, J. (1997). Mobility performance with retinitis pigmentosa. *Clinical and experimental optometry*, *80*(1), 1-12.

Colenbrander, A. (2002). Visual standards: aspects and ranges of vision loss with emphasis on population surveys. *Report for the International Council of Ophthalmology*, *2*, 1-33.

Chen, J., Tong, Y., Gray, W., & Ji, Q. (2008). A robust 3D eye gaze tracking system using noise reduction. In *Proceedings of the 2008 symposium on Eye tracking research & applications*(pp. 189-196). ACM.

Coeckelbergh, T. R. M. (2002). *Effect of compensatory viewing strategies on practical fitness to drive in subjects with visual field defects caused by ocular pathology*. University Library Groningen

de Haan, G. A., Heutink, J., Melis-Dankers, B. J., Tucha, O., & Brouwer, W. H. (2014). Spontaneous recovery and treatment effects in patients with homonymous visual field defects: a meta-analysis of existing literature in terms of the ICF framework. *Survey of ophthalmology*, *59*(1), 77-96.

de Haan, G. A., Heutink, J., Melis-Dankers, B. J., Brouwer, W. H., & Tucha, O. (2015). Difficulties in daily life reported by patients with homonymous visual field defects. *Journal of Neuro-Ophthalmology*, *35*(3), 259-264.

de Haan, G. A., Melis-Dankers, B. J., Brouwer, W. H., Tucha, O., & Heutink, J. (2016). The effects of compensatory scanning training on mobility in patients with homonymous visual field defects: further support, predictive variables and follow-up. *PloS one*, *11*(12), e0166310.

Duchowski, A. T. (2007). Eye tracking methodology. *Theory and practice*, *328*.

Eckstein, M. K., Guerra-Carrillo, B., Singley, A. T. M., & Bunge, S. A. (2017). Beyond eye gaze: What else can eyetracking reveal about cognition and cognitive development?. *Developmental cognitive neuroscience*, *25*, 69-91.

Edwards, J. D., Ross, L. A., Wadley, V. G., Clay, O. J., Crowe, M., Roenker, D. L., & Ball, K. K. (2006). The useful field of view test: normative data for older adults. *Archives of Clinical Neuropsychology*, *21*(4), 275-286.

Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in cognitive sciences*, *9*(4), 188-194.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. OUP Oxford.

Kwon, M., Huisingh, C., Rhodes, L. A., McGwin Jr, G., Wood, J. M., & Owsley, C. (2016). Association between glaucoma and at–fault motor vehicle collision involvement among older drivers: a population-based study. *Ophthalmology*, *123*(1), 109-116.

Lee, S. S. Y., Black, A. A., Lacherez, P., & Wood, J. M. (2016). Eye movements and road hazard detection: effects of blur and distractors. *Optometry and vision science*, *93*(9), 1137-1146.

Lee, S. S. Y., Black, A. A., & Wood, J. M. (2017). Effect of glaucoma on eye movement patterns and laboratory-based hazard detection ability. *PloS one*, *12*(6), e0178876.

MacInnes, J. J., Iqbal, S., Pearson, J., & Johnson, E. N. (2018). Wearable Eye-tracking for Research: Automated dynamic gaze mapping and accuracy/precision comparisons across devices. *bioRxiv*, 299925.

Nilsson, M., & Nivre, J. (2011). A survival analysis of fixation times in reading. In *Proceedings of the 2nd Workshop on Cognitive Modeling and Computational Linguistics* (pp. 107-115). Association for Computational Linguistics.

Olsen, A. (2012). The Tobii I-VT fixation filter. *Tobii Technology*.

Petzold, A., & Plant, G. T. (2005). Central and paracentral visual field defects and driving abilities. *Ophthalmologica*, *219*(4), 191-201.

Pambakian, A. L. M., Wooding, D. S., Patel, N., Morland, A. B., Kennard, C., & Mannan, S. K. (2000). Scanning the visual world: a study of patients with homonymous hemianopia. *Journal of Neurology, Neurosurgery & Psychiatry*, *69*(6), 751-759.

Papageorgiou, E., Hardiess, G., Schaeffel, F., Wiethoelter, H., Karnath, H. O., Mallot, H., ... & Schiefer, U. (2007). Assessment of vision-related quality of life in patients with homonymous visual field defects. *Graefe's Archive for Clinical and Experimental Ophthalmology*, *245*(12), 1749-1758.

Pollock, A., Hazelton, C., Henderson, C. A., Angilley, J., Dhillon, B., Langhorne, P., ... & Shahani, U. (2011). Interventions for visual field defects in patients with stroke. *Cochrane Database of Systematic Reviews*, (10).

Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. In *Proceedings of the 2000 symposium on Eye tracking research & applications*(pp. 71-78). ACM.

Warren, M. (2009). Pilot study on activities of daily living limitations in adults with hemianopsia. *The American Journal of Occupational Therapy*, *63*(5), 626. Chicago

Williams, R. A., Brody, B. L., Thomas, R. G., Kaplan, R. M., & Brown, S. I. (1998). The psychosocial impact of macular degeneration. *Archives of Ophthalmology*, *116*(4), 514-520.

Zihl, J. (1995). Visual scanning behavior in patients with homonymous hemianopia. *Neuropsychologia*, *33*(3), 287-303.