Investigating the Central Fixation Bias

Untersuchung zur zentralen Fixationstendenz

by

Lisa Schwetlick

A thesis submitted to the
University of Potsdam, Germany
in partial fulfillment of the requirements for the degree of
BACHELOR OF SCIENCE IN PSYCHOLOGY

Supervisors

Prof. Dr. Ralf Engbert
Dr. Hans Trukenbrod
Lars Rothkegel

Department of Biopsychology
University of Potsdam, Germany

June 2016

Abstract

The central fixation bias is an effect that is ubiquitous in scene viewing experiments. Human observers display a tendency to fixate the center of an image initially and longer than the periphery. This can be observed irrespective of image feature distribution, image characteristics, as well as screen- and fixation marker positions. As it is likely to cover up less pronounced effects it is sometimes necessary to control it. To gain a better understanding of the central fixation bias we propose an experiment to dissociate between two possible explanations for this phenomenon. The gist extraction approach proposes that the image center is optimal for gaining a basic understanding of the image. The alternative explanation is that the sudden luminance change at image presentation attracts the eyes to its center. The experiment examined the effects of two different types of image onset and the presence or absence of a peripheral preview on the central fixation bias. We conducted a pilot study to assess the suitability of this method and were able to successfully implement both the image onset type and and preview factors. However, more data is needed to draw reliable conclusions about the causes for the central fixation bias.

Zusammenfassung

In der Blickbewegungsforschung wird häufig beobachtet, dass Versuchspersonen die Mitte eines Bilds zuerst und häufiger fixieren als die Peripherie. Diese zentrale Fixationstendenz ist unabhängig von Bildeigenschaften sowie Fixationskreuz- und Bildschirmposition.

Eine mögliche Erklärung für dieses Phänomen ist, dass die Mitte des Bildes die optimale Position ist, um grundlegende Informationen über das Bild zu extrahieren. Dementsprechend sollten Betrachter, wenn diese Information ihnen bereits vorliegt, eine weniger ausgeprägte zentrale Fixationstendenz zeigen.

Die alternative Erklärung ist, dass in Blickbewegungsexperimenten Bilder häufig in Zusammenhang mit einer plötzlichen starken Luminanzänderung des Bildschirms auftauchen. Solche Luminanzveränderungen ziehen den Blick typischerweise an. Wenn dies der Fall ist, dann sollte der Effekt verringert werden, wenn das Bild nicht mit einem Mal aufblitzt.

In einem Pilotexperiment haben wir diese beiden Erklärungsansätze gegeneinander aufgewogen. Bilder wurden entweder mit einem Mal präsentiert oder langsam eingeblendet, um den Effekt des Aufblitzens zu messen.

Um den Einfluss vorheriger Bildinformation zu messen, konnten die Bilder in vier Positionen präsentiert werden; entweder sie waren schon zu sehen bevor sie direkt angeschaut wurden oder sie wurden erst zum Zeitpunkt eingeblendet. Durch einen peripháren Eindruck des Bildes können die Versuchspersonen bereits erste Informationen über das Bild gewinnen. Aus diesen Faktoren ergibt sich ein 2x2 Experimentaldesign.

Im Rahmen dieser Arbeit war zu testen, ob die vorgeschlagene Methode geeignet ist, um die zentrale Fixationstendenz zu untersuchen. Die Umsetzung des Experiments war grundsätzlich erfolgreich. Die Versuchspersonen konnten dem Quadrantendesign folgen und wichen nur selten von der vorgeschriebenen Reihenfolge ab. Auch das Einblenden der Bilder als Umsetzung des Gegenteils des Aufblitzens scheint legitim zu sein. Dabei sollte eventuell überprüft werden, ob es in allen Konditionen dieses Experiments verwendbar ist.

Rückschlüsse über die Ursachen der zentralen Fixationstendenz können erst mithilfe einer größeren Stichprobe reliabel gezogen werden.

Acknowledgements

Firstly, I would like to thank my supervisors Lars Rothkegel, Dr. Hans Trukenbrod and Prof. Dr. Ralf Engbert for their continued support, valuable guidance and encouragement during the writing of this thesis. I give special thanks to Lars Rothkegel, who was most closely involved in this project. Thank you for always taking time to answer my questions, share expertise and give feedback. My gratitude also goes out to the staff of the University of Potsdam Eyelab for meticulously conducting the experiment. Lastly, I am very grateful to the invaluable people who read drafts of this thesis and offered constructive criticism.

Contents

1 Introduction

Our sense of sight provides us with a seemingly seamless view of the world. In fact, we see is recreated in the brain from a multitude of snapshots, focused on different areas in the scene. To focus on and center different elements of interest in the visual field, the eyes move in a succession of jumps, or saccades, around the scene. These saccades are guided by a variety of cognitive processes. The link between these processes and eye movements provide a window into the brain and allows us to draw conclusions about the mechanisms of attention, memory and perception.

1.1 Motivation and Problem Statement

The *central fixation bias* is encountered sooner or later in most scene viewing experiments. It is remarkable how it is invariably found that subjects fixate the center of an image initially, longer, and more frequently than the edges. The central fixation bias persists throughout changes in image characteristics, fixation marker location, and screen location. Researchers often attempt to prevent or compensate for this tendency because it is likely to cover up less pronounced effects.

Recent research has shown that the central fixation bias might be a consequence of initial gist extraction from the image center. Centering the middle of an image in the visual field allows the rest of the image to be peripherally processed, in ways that can shape subsequent viewing behavior.

An alternative explanation is that scene viewing experiments usually encompass a sudden luminance change when the image is first presented. Such luminance flashes are known to attract the eyes. It is possible that the observed bias towards the center of the image appears because the eye is attracted to the area where the change in the visual field is largest, i.e. the center.

1.2 Proposed Approach

In this thesis we will attempt to dissociate between these two explanations for the central fixation bias: gist extraction from the center of the image and capture of the visual attention caused by sudden image onset and the related luminance change. We therefore propose an eye tracking experiment that measures the eye movements of subjects in response to images presented in 4 conditions of a 2x2 factorial design.

The first factor is the presence or absence of sudden luminance change. Its absence will be implemented as a gradual fade effect; its presence through the traditional sudden presentation.

The second factor is previous gist extraction from, or ignorance of, the image. This can be implemented by presenting four images simultaneously or consecutively. Simultaneous presentation allows the viewer to extract image information from the periphery of their current visual field, giving them an idea of the image before they view it directly. By contrast, consecutive presentation leaves the viewer naive in terms of image information until they actively examine it. The results will be evaluated using distance to center measurements.

1.3 Structure

This thesis we will firstly examine the central fixation bias in more detail and discuss the merit of different explanations for it in the Theoretical Background section (sec. 2). Section 3 (Method) will provide detail on the conditions and experimental procedure. The results of the data analysis are described in section 4 (Results) and discussed in section 5 (Discussion).

1.4 Research Questions

In order to define the scope and focus of this thesis, the following research questions (RQ) were defined:

RQ1 Quadrants: Is the quadrant design and the associated peripheral preview adequate to implement the condition of gist extraction?

RQ2 Fade: Is the fade manipulation a legitimate operationalization of "no sudden luminance change"?

RQ3 Causes: Can this experimental setup provide insight into the causes for the central fixation bias?

2 Theoretical Background

Every day we are exposed to a vast number of sensory stimuli. It is impossible to attend to all of these equally and simultaneously, as our brain's processing resources are limited. The brain compensates for this by using efficient methods of gathering important information while disregarding the rest. This mechanism of allocating processing resources is called attention.

The field of attention research is vast. One notable and relatable example is the "Cocktail party effect", i.e. the question of how we can focus our auditory attention on one stream of information in the presence of multiple simultaneous conversations, as one would at a busy cocktail party (Arons, 1992). When questioned about the unattended auditory streams, subjects were usually unable to answer questions about the content (Broadbent, 1954; Cherry, 1953), but were sometimes able to react to semantically interesting information in an unattended channel (e.g. their own name)(Wood & Cowan, 1995).

Another similar effect in the field of visual attention is known as "inattentional blindness". Simons and Chabris (1999) demonstrated this memorably in their "Invisible Gorilla Experiment". Participants watched two teams, one dressed in black, the other in white, throwing basketballs and were asked to count the amount of passes made by the white team. Most did not notice a man dressed as a black gorilla walking through the middle of the picture. Their attention was so focused on the white players, that even something as absurd as a gorilla was completely disregarded. This shows that attention is a very selective process and illustrates how we are unable to keep track of all the sensory input we get.

The visual system is further limited in what it can perceive. Although we typically experience the world as seamless, high color and in focus, only a very small part of our visual field actually produces a high resolution image. The *fovea*, in the center of the retina, is several times more sensitive than the periphery (Bear, Conners, & Paradiso, 2007). In order to compensate for this, humans move their eyes over scenes in a succession of short ballistic movements called *saccades* during which the eye is essentially blind. These are alternated with *fixations* of 150-250 ms (Salthouse & Ellis, 1980), where information is registered. Using these snapshots of a scene, the brain creates the illusion of smooth, coherent visual input.

Two separate forms of attention can be distinguished: overt attention, when the focus of the visual field and the focus of attention align, and covert attention, when they do not. (Posner, 1980). In overt attention the eyes are directed at the element that is being attended. As the center of the retina can perceive much more detail than the periphery, it seems natural to center elements of interest in the visual field for examination. This instinctual tendency is very useful for exploring visual attention; it allows us to observe the locus of attention moving across an image, usually with the help of an eye tracking device.

Covert attention, as a shift in attention independent of the eye's position, is harder to detect. However, recent research by Engbert and Kliegl (2003) suggests that *microsaccades*, tiny jitters of the eye during fixations, can imply the direction of the covert attention.

2.1 Guidance Principles

It is understood that covert shifts in attention usually precede the actual movement of the eye and thus the onset of overt attention (Hoffman, 1998). Hoffman states that "the relationship between attention and eye movements is one of partial interdependence. Attention is free to move [independently], but the eyes require visual attention to precede them to their goal". This is the case regardless of whether movement is triggered by internal guidance or external factors (Yantis & Jonides, 1996). We can therefore assume that at some point during a fixation, the location of the next fixation is chosen from the periphery of the current visual field, according to various guidance principles.

2.1.1 Top down

Top down mechanisms are dependent on the viewer, their knowledge, expectations and intentions. In his 1967 work "Eye Movements and Vision", Yarbus describes an experiment in which subjects were asked to examine a scene with varying task instructions. He found visible differences in the eyes' paths depending on what the participant had been asked to do (see fig. 1).

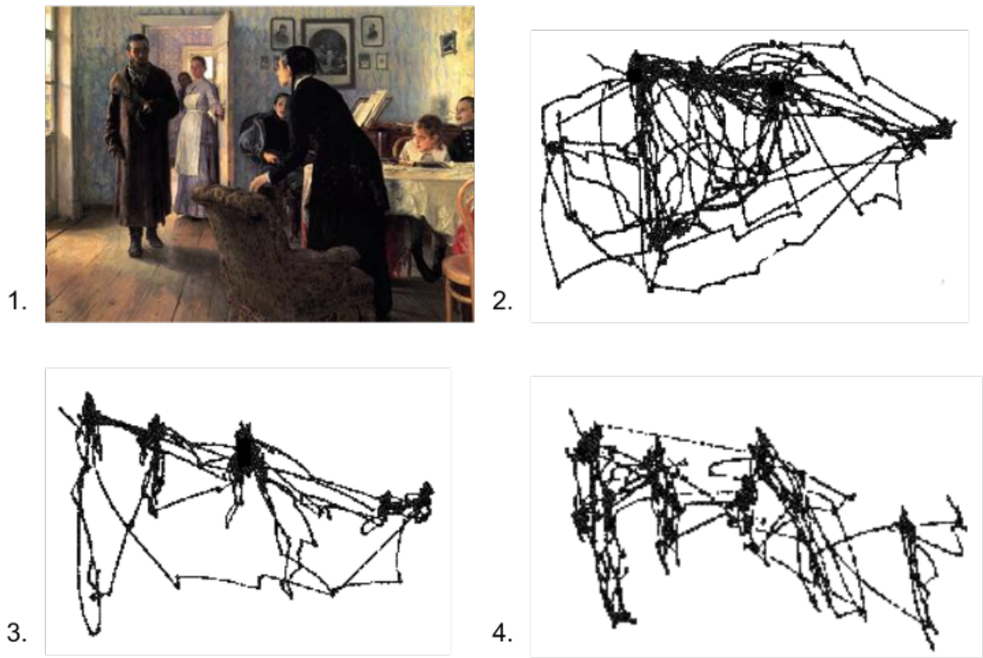

Figure 1: (1) shows the original picture that was shown to the participants in Yarbus' experiment, "The Unexpected Visitor" by Ilya Repin. The other images show the scan paths that the subject's eyes took across the image, depending on the task they had been given: (2) free viewing, (3) to guess the ages of the people in the picture and (4) to remember the clothes worn.

Further research also found that areas of a picture that were independently rated as informative were more likely to be fixated than uninformative ones (Mackworth & Morandi, 1967). More recent studies have tried to dissociate between image feature influences and top down influences. These showed that unexpected elements in scenes, such as an octopus on a farm, were fixated longer and more frequently than expected elements, such as a tractor (Henderson, Weeks, Phillip A., & Hollingworth, 1999; Loftus & Mackworth, 1978). This lends credibility to the hypothesis that conscious thought and expectations have an effect on our viewing behavior. It is generally assumed that top down processes mainly influence later viewing behavior, while the first fixations are guided by the properties inherent in images, by bottom up processes (Torralba, Oliva, Castelhano, & Henderson, 2006).

2.1.2 Bottom up

Bottom up mechanisms are dependent on the characteristics of the image itself. They are task independent (Itti, Koch, & Niebur, 1998) and rely on features like luminance, local contrast, edge density, and regions of high spatial frequency content (Mannan, Wooding, & Ruddock, 1996) to guide the eye. When modeling viewing behavior with computers, image feature information can be used to create *saliency maps*, which, according to Itti et al. (1998), "topographically [code] for conspicuity over the entire visual scene". These computed maps can then be compared to human behavior to verify their accuracy.

Studies examining the effect of different image features on viewing behavior come to different conclusions regarding their importance (Reinagel & Zador, 1999; Rentschler, Hauske, Schill, Zetzsche, & Krieger, 2000). Parkhurst and Niebur (2003) showed that there is a difference in terms of image statistics between fixated points and other available points in the image.

2.1.3 Systematic

The last, least cited, category of influences on eye guidance are systematic tendencies. These are tendencies that arise neither from the viewer's condition, nor directly from image features. The position of the eyes in the head and the way that muscles function, laboratory environments, as well as the way photos of natural scenes are taken, all severely limit *how* we view scenes. These systematic tendencies are very influential and could "explain a high proportion of the variance in where people look", according to Foulsham and Kingstone (2012).

An example for this is *saccadic momentum*, the phenomenon that executing a saccade in the opposite direction to the previous saccade will result in a longer fixation duration and that saccades ion the same sirection are more likely (Smith & Henderson, 2009). Another important systematic tendency is the *central fixation bias*.

2.2 Central Fixation Bias

The central fixation bias (CFB) is a systematic tendency of observers to look at the center of an image more frequently than at the edges (see fig. 2)(Tatler, 2007).

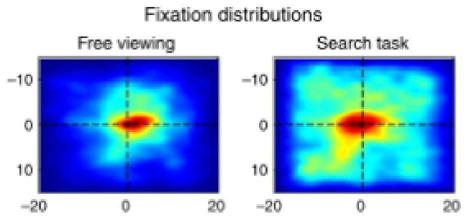

Figure 2: Typical distribution of fixations in a scene viewing experiment (Tatler, 2007). The center of the image is fixated more frequently than the edges, regardless of the task.

This was observed by Buswell as early as 1935 and is well documented in scene viewing literature. Although many models, computational or otherwise, try to predict where fixations will be made using advanced image statistics, even the best model performs only slightly better than a simple model based of the distance to the center(Judd, Ehinger, Durand, & Torralba, 2009). Being so fundamental in nature, the central fixation bias becomes a point of discussion in many scene viewing experiments and, when researching image saliency, often has to be compensated for, as it otherwise overshadows other relevant data. In order for such a compensation or prevention to take place, we need to understand the underlying principles. The following section will discuss some of the possible as well as the rejected explanations for the central fixation bias.

2.2.1 Image features

The first and most commonly assumed explanation for the CFB is the photographer's bias: people taking pictures will usually make sure that the subject of their photograph is centered and in focus. Therefore, if we assume a correlation between image features and fixations (Mackworth & Morandi, 1967) and assume that the features are cluttered in the center, there will be a central fixation bias. Several recent studies have tried to prevent this effect by choosing images that had no evident central feature bias. Parkhurst and Niebur (2003) found that, although it did not disappear, the CFB did tend to shift in the direction of the feature distribution. Tatler (2007) also found a strong, persevering CFB, despite having laterally biased images.

2.2.2 Motor biases

Another popular explanation is that scene viewing experiments typically precede each image with a centrally placed fixation marker. This starting point and the fact that we are disposed to make small amplitude saccades (Bahill, Adler, & Stark, 1975; Tatler, Baddeley, & Vincent, 2006), may lead to the eyes simply never leaving the surroundings of the fixation marker, and produce the observed effect. Tatler (2007) tested this hypothesis by postulating that if the motor bias was the reason for the CFB then a random walk model, in which each step had the characteristics of a human saccade, should produce a similar bias as the eye movements. For both randomized and central starting positions the random walk model failed to account for a bias as strong as the CFB. Irrespective of the location of the fixation marker, human subjects always made their initial fixations in the center of the image. After these first two traditional explanations failed to account for his data, Tatler (2007) suggested the three following three hypotheses: centering the eye in its orbit, convenient exploration starting point and gist extraction.

2.2.3 Centering the eye in its orbit

Due to the physiology of the eye, it is possible that looking straight ahead is the most relaxed, comfortable state for the eye to be in. As participants in psychological experiments are generally seated in front of the the screen so that looking straight ahead would also explain the CFB. However, as shown by Vitu, Kapoula, Lancelin, and Lavigne (2004), positioning the screen so that its center does not align with the straight ahead position of the eyes does not cause the CFB to disappear.

2.2.4 Gist extraction

The tendency to fixate the center of the screen first might also be helpful on a semantic level. It allows the visual field to encompass as much of the scene as possible. Although this information will not be detailed, it may allow gist extraction and orienting in the scene (Torralba et al., 2006). Rough gist extraction from an image can happen within 50 ms (Bacon-Macé, Macé, Fabre-Thorpe, & Thorpe, 2005). A 75 ms preview of an image can lead to a greater proficiency in searching the image (Võ & Henderson, 2010), suggesting that such a preview helps with orientating in the scene.

A recent study in our lab tested this hypothesis with the following experiment (Rothkegel, Trukenbrod, Schütt, Wichmann, & Engbert, 2016). Participants were shown images of scenes on the computer screen. They were given a fixed starting point somewhere in the image by means of a fixation marker and were asked to keep their eyes on it, until it disappeared. The image was presented underneath the marker, effectively forcing participants to fixate one location in the image for a certain amount of time (moving away from it would cause the image to disappear). After this phase, participants were allowed to freely explore the scene. If gist extraction was indeed the cause for the CFB, participants could be forced to extract the gist from a non-central predetermined position, and we would expect to see reduced central tendencies once the free viewing phase started. As shown in fig. 3, the CFB was significantly reduced even when participants were shown previews of only 0.125s. These results seem to lend a lot of credence to the proposed hypothesis that the CFB is an effect of gist extraction.

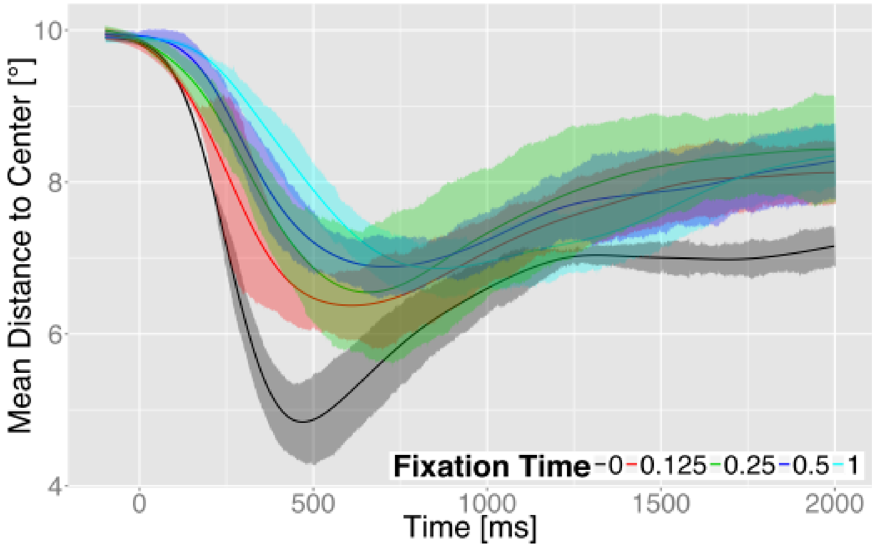

Figure 3: Data from a central bias experiment in our lab by Rothkegel et al. (2016). Preview fixation times of as little 0.125 s (red) can significantly reduce the CFB compared to the no preview (black) condition. This is particularly evident in the first 800 ms.

2.2.5 Convenient for starting the exploration

The center of the screen may be a strategically convenient location to begin exploration of the scene. Starting in the center minimizes the distance to every other point in the scene. However, this hypothesis can not provide an explanation for the absence of the CFB in the experiment described above. If the center is just a convenient starting location with no semantic benefit, this effect should not be negated by a preview.

2.2.6 Luminance change

In most eye tracking experiments images are presented onto the screen all at once, producing a sudden onset luminance change, or *flash*. It has been shown in multiple experiments that luminance changes and movement attract the eyes (Theeuwes, 1995; Theeuwes, Kramer, Hahn, & Irwin, 1998; Irwin, Colcombe, Kramer, & Hahn, 2000). In the gist extraction experiment described above, this flash effect is not present, since the eye position at the moment of picture onset is fixed. While the setup allowed fixed position previews, it also resulted in the flash that is normally found in such experiments being omitted. It is not clear therefore, whether the missing image onset flash or the preview caused the observed results.

2.3 Experiment Proposition

In this paper we propose an experiment that will dissociate between the preview and flash explanations for the results found in the experiment by Rothkegel et al. (2016) described above.

3 Method

The experiment followed a 2x2 design with the following conditions:

1. flash and preview
2. no flash and preview
3. flash and no preview
4. no flash and no preview

Additionally, a fifth, supplementary control condition was added to account for the rarely used fade manipulation we used (see Operationalization, sec. 3.4.1 below).

The experiment was programmed in *MATLAB R2015a* (2015) using the PsychToolbox toolkit (Brainard, 1997). It was divided five blocks, one per condition, of 20 trials. During each trial four images were presented according to one of the five conditions, and examined successively. In the pilot study the blocks appeared in the same order for all participants.

3.1 Stimuli

The experiment was conducted with a set of 400 natural scene images. They were presented at 467x467 px with a resolution of 96 dpi in each direction. This was the largest possible resolution for presenting four images on the given monitor. At a viewing distance of 60 cm each image subtended 11.76° of visual angle.

The square format of the images ensured an equal distance when moving from one image to the next vertically and horizontally. It furthermore had the advantage of allowing us to crop the original pictures in such a way that would minimize the photographer's bias of placing salient features in the center. Each stimulus consisted of four equidistant images. The images appeared consistently in the same position over all trials and participants to reduce position-induced variance. The stimulus material was presented on a monitor with the following properties.

- resolution of 1280 x 1024 pixels
- refresh rate of 60Hz

3.2 Participants

The eye movements of five human participants were recorded in the context of a pilot study. The group consisted of 1 male and 4 females with normal or corrected to normal vision.

3.3 Eye Movement Recording

Eye movements were recorded at a sampling rate of 1000Hz using an Eyelink 1000 video-based eyetracker (SR-Research, Osgoode/ON,Canada). Only the right eye was tracked. A 9-point target grid was used for calibrating the eye position. The calibration was then validated by a second 9-point target grid. If the validation diverged sufficiently from the calibration values, the calibration was re-initiated. After the initial calibration this procedure was repeated every 10 to 14 trials.

3.4 Procedure

Participants were instructed to seat themselves comfortably at a viewing distance of 60 cm from the screen with their heads positioned on a chin rest. They were told that they would be participating in a memory experiment. This was done in order to encourage the participants to remain attentive throughout the experiment. Subjects were also given the constraint of looking at the pictures in a prescribed order (top left, top right, bottom right, bottom left) so the distance that their eyes had to move from one image to the next was held constant (as opposed to the traditional left to right behavior, where the jump from the top right to the bottom left image would have been longer than between any other images). Participants were told to move on to the next image only after the current one had disappeared. The instructor was able to see the eye movements in real time and could point out incorrect viewing behavior. Participants completed 5 blocks in order, each consisting of 20 trials. Each trial started with a fixation cross at the center of the screen followed by the presentation of four images according to one of the following conditions. Every 5-7 trials a memory prompt was shown in which participants were asked to determine whether they had seen a specific image since the last prompt (see fig. 4).

Figure 4: Screenshot of the memory prompt. The title translates to "Did you see this picture in the last block?", where the block refers to the last 5-7 presentations since the last prompt.

3.4.1 Operationalization

The *preview* conditions were implemented by presenting four images on the screen in each trial, allowing a peripheral view and gist extraction of all images when focusing on a single one. *No preview* conditions showed only one image at a time.

The onset factor was implemented by presenting images using a fade effect for *no flash* conditions and all at once for *flash* conditions. We chose to present images over a black background in order to maximize the luminance change and generate a stronger flash. The fade effect has the disadvantage of being a very rarely employed technique in scene viewing experiments. This made it unclear which consequences it might have. In an effort to keep the *flash* and *no flash* conditions as similar and as comparable as possible, we introduced a *norm fade* phase. This phase describes a short fade from 0-10% opacity that occurs in all of the main conditions. It is short enough to not disrupt the flash effect, but introduces the fade into every condition. As another precaution to make the results more comparable we added a fifth condition to the 2x2 design, which omits any kind of fade effect. If the results with and without the fade are similar, we can safely assume that the fading mechanic did not introduce any large unexpected artifacts to the experiment. In accordance with the proposed hypotheses that sudden image onset and/or gist extraction from the image center cause the central fixation bias, the *no preview flash* condition should produce the strongest bias. Correspondingly, we expected the *preview no flash* and *control preview no flash* conditions to cause the weakest biases.

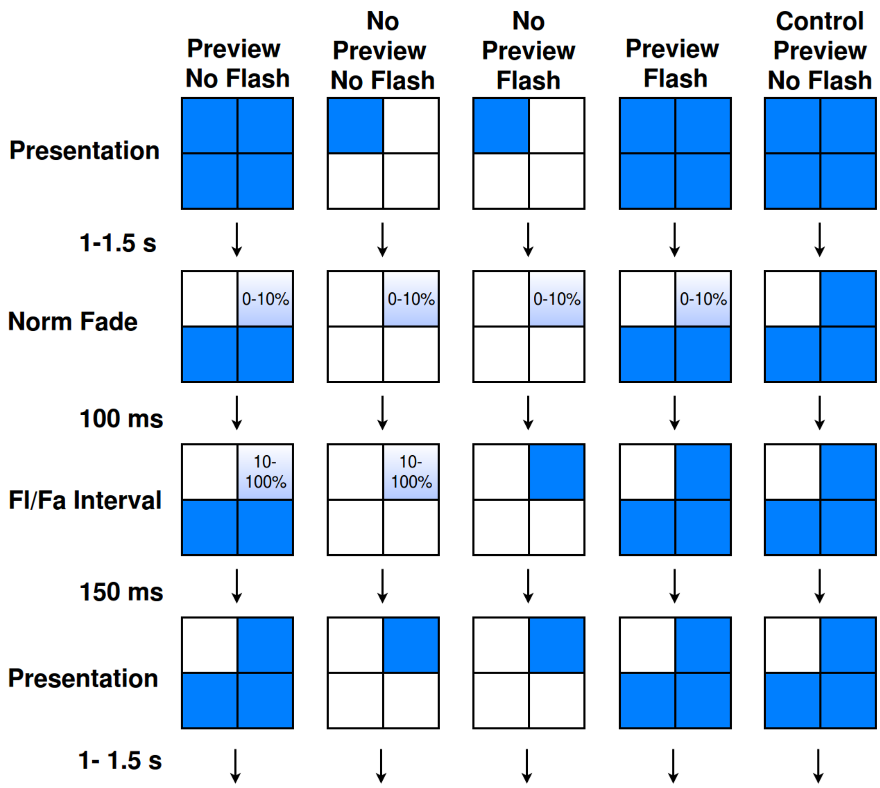

Figure 5: This image shows how each condition changes over time. Blue squares represent an image that was shown, white squares represent images that were not shown. The percentages inside the lighter blue squares indicate a fade taking place, from the first to the second percentage of opacity.

As per instruction, the participants looked at images successively, moving to the next image only after the current one disappeared. In order to avoid anticipatory, premature saccades to the next image the presentation time varied randomly between 1 and 1.5 s along a uniform distribution.

Fig. 5 shows the exact experimental procedure, including all the described phases.

3.4.2 Preview & flash

As shown in fig. 6 all four images are shown at the beginning of the trial. After the presentation duration of 1-1.5 s the first image disappears. The second image also disappears but immediately begins to fade back in, up to a opacity of 10% in 100 ms (norm fade). Next, the second image is presented at full opacity, constituting the luminosity change of the flash, and is subsequently presented for 1-1.5 s. This process is repeated for images 2, 3 and 4.

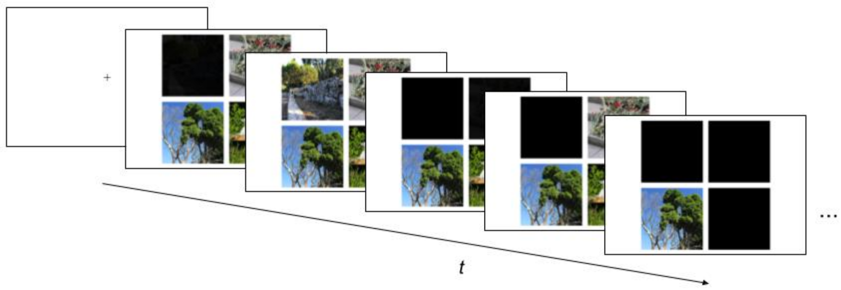

Figure 6: Timeline of the *preview & flash* condition

3.4.3 Preview & no flash

As before, all images are present at the beginning of the trial to allow preview. When the first image disappears after 1-1.5 s the second image is faded in to 10% opacity (norm fade). In this condition, however, the first fade is followed by a second fade of 150 ms up to 100% opacity, circumventing the abrupt change in luminosity. Finally the second image is presented like the first one for 1-1.5 s. This process is repeated for images 2, 3 and 4 (fig. 7). This condition excludes both factors, central gist extraction and sudden onset, which we hypothesize to cause to central fixation bias. We, therefore, expect it to cause the least strong bias.

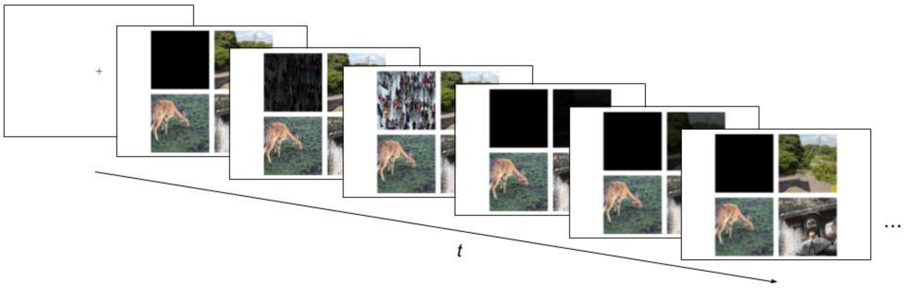

Figure 7: Timeline of the *preview & no flash* condition

3.4.4 No preview & flash

In the no preview conditions only one image at a time is visible. The first image is presented on its own for 1-1.5 s. When it disappears, the second image performs the norm fade to 10% opacity and is then immediately presented at full opacity (flash). This process is repeated for images 2, 3 and 4 (fig. 8). As it includes both factors that may contribute to the central fixation bias, we expected this condition to evoke the strongest bias.

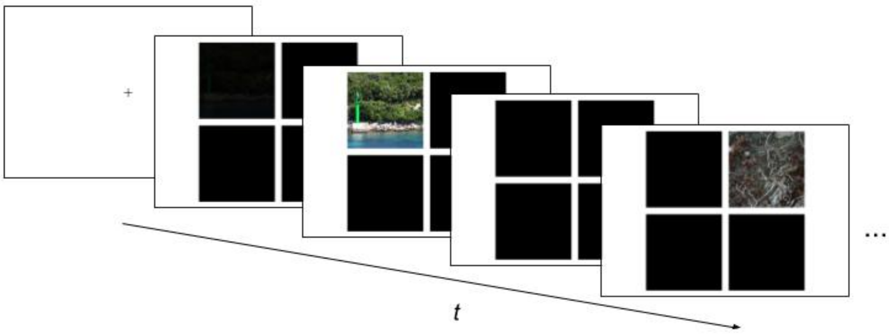

Figure 8: Timeline of the *no preview & flash* condition

3.4.5 No preview & no flash

Again, only one image is visible at a time. The first image is presented for 1-1.5 s. Then the second image performs the norm fade and is then faded in to 100% opacity over 150 ms. This process is repeated for images 2, 3 and 4 (fig. 9).

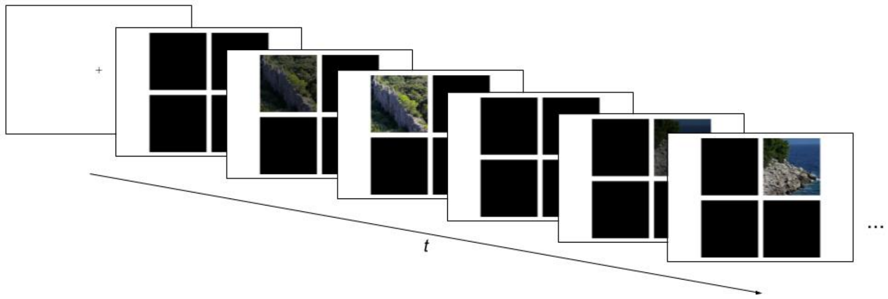

Figure 9: Timeline of the *no preview & no flash* condition

3.4.6 Control: preview & no flash

All images are present at the start (preview). The first image disappears after 1-1.5 s. The second image is not faded in, but simply stays visible (no flash). This condition was added to control for the fact that most experiments do not use any fades, and to make the results more comparable (fig. 10).

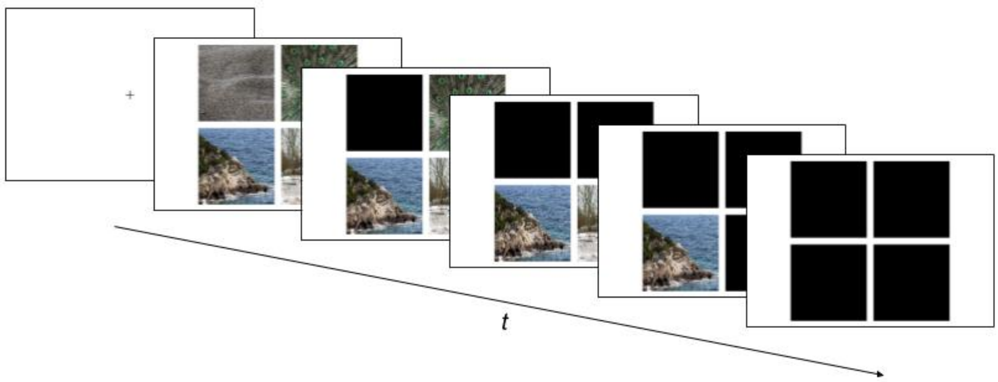

Figure 10: Timeline of the *control preview & no flash* condition

4 Results

In a pilot study with 5 subjects we collected the following results to show that the described experiment is suited to research causes for the central fixation bias. In the scope of this thesis we will focus on whether the experiment was conducted successfully.

In the evaluation we considered only fixations with durations of more than 50 ms. Fixations outside the four quadrants were disregarded as we were interested only in the fixations inside the image quadrants. The first step is to determine whether the viewing behavior elicited by the experiment is compatible with the expectations.

4.1 Viewing Order Compliance

Although the instructions specifically asked subjects to closely follow the correct viewing order, not all fixations landed in the intended quadrants. Fig. 11 shows the absolute amount of times each subject jumped ahead of the prescribed order in each condition. There was a significant difference between subjects. Subjects 1 and 2 in particular made a large number of anticipatory saccades. The maximum possible amount of anticipatory saccades per condition was 60 (80 images presented per condition in total, 20 of which can be disregarded because they were presented in position 1). In the *preview flash* condition subject 2 showed anticipatory saccades in a third of all images.

Conditions the *preview* and *control* conditions had significantly more anticipatory saccades. This was to be expected, because in the other two conditions no images were presented before they were supposed to be looked at.

As fig. 12 shows, there was no significant amount of returning to previous images for any of the main conditions. The *control* condition elicited a far larger amount of backward jumps. Again, the maximum possible amount per subject and condition was 60, meaning that subject 3 showed backwards jumps in almost half the images in the control condition.

Fig. 13 shows that all images were fixated roughly the same amount of

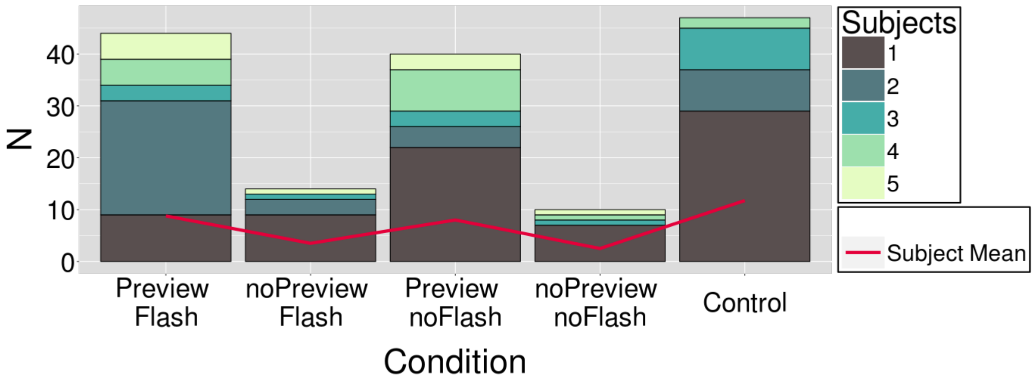

Figure 11: Amount of times subjects moved to the next location before they were supposed to (anticipatory saccades) across conditions. The maximum possible amount per subject was 60.

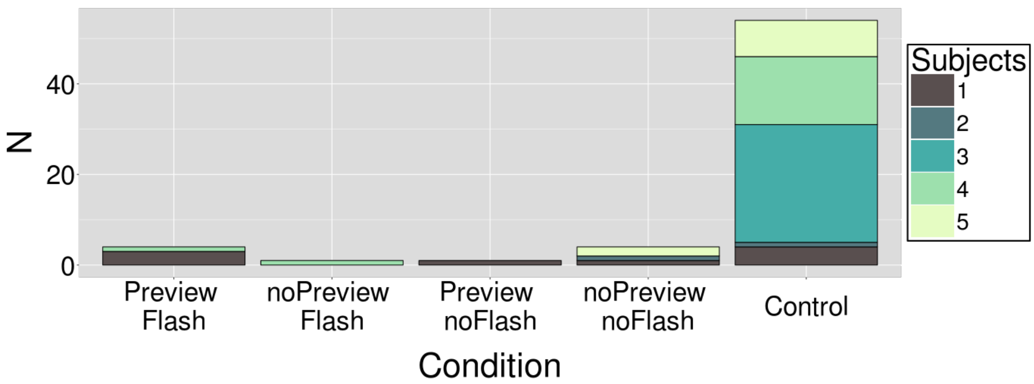

Figure 12: Amount of times subjects returned to previous locations across conditions. The maximum possible amount per subject was 60.

times, regardless of the participant and the position within the trial. Position 4 elicited slightly fewer fixations on average. This might be a consequence of the fact that the eyes move with a slight delay to the actual presentation time. The recording of eye movements stopped after the allocated time in position four was over, not accounting for this delay.

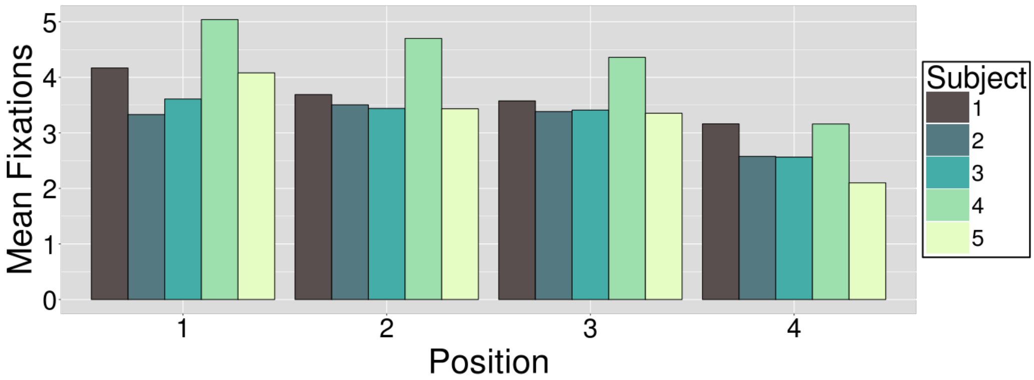

Figure 13: Amount of fixations for each subject in each of the four image positions in the trials.

4.2 CFB Phenomenon

For each fixation the distance to the center of the corresponding image was calculated. In these visualizations we excluded data from position 1, as this position would not have included a preview and is qualitatively different from data gathered from positions 2 to 4. Using this method, we were able to find a central fixation bias in the collected pilot data. Fig. 16 shows the distribution of fixations over all trials and positions. Participants initially fixated the center of the image before moving to the periphery, as expected.

By taking a closer look at the distance to the center over the fixations we found the following phenomenon. Fig. 15 shows that if all fixations over 50 ms are included, the second fixation is actually closer to the center than the first. However, if we exclude very short saccades (only including saccades over 150 ms or 200 ms in length, respectively), the curve, on average, conforms more to our expectations, with the first fixation being the closest to the center and the following ones increasingly further away. Viewing this figure it is important to note that deviations from this tendency in the late fixations may be unreliable because there are only few data points. The subjects usually did not have time to make more than 4 or 5 fixations per image, with very short fixations included.

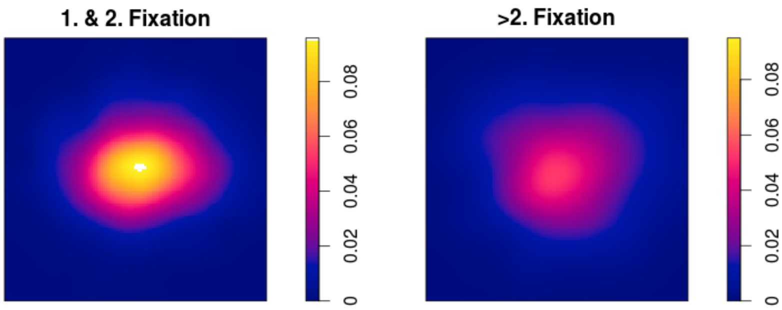

Figure 14: Distribution of all fixations, in all positions except position 1. Much higher density of fixations can be observed in the center of the image, especially in the first two fixations. This tendency dissipated somewhat in later fixations.

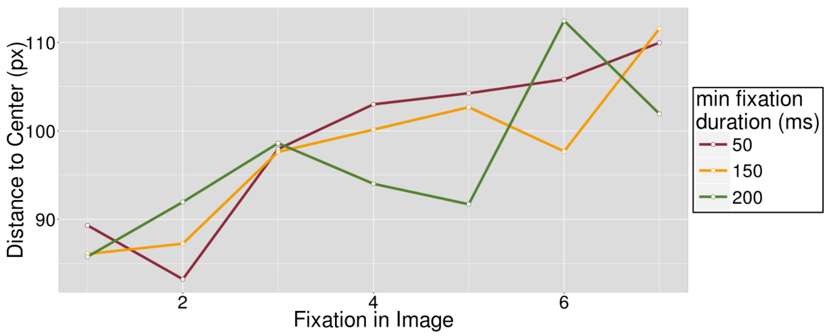

Figure 15: Average distance to the center for all fixations longer than 50, 100 or 200 ms. When considering very short fixations, the second fixation is closest to the center. For longer fixations, the first fixation is closest.

Another observation was that the CFB tended to be distorted in the direction of the previous image. Fig. 16 *a)* shows the CFB for the first two fixations in each position, over all trials. In the first position (i.e. the upper left), subjects arrived from the fixation cross in the middle, giving the distribution a tendency toward the bottom right. The distributions in positions 2 through 4 all tend towards the preceding position. In the second fixation, *b)*, this distortion is no longer visible.

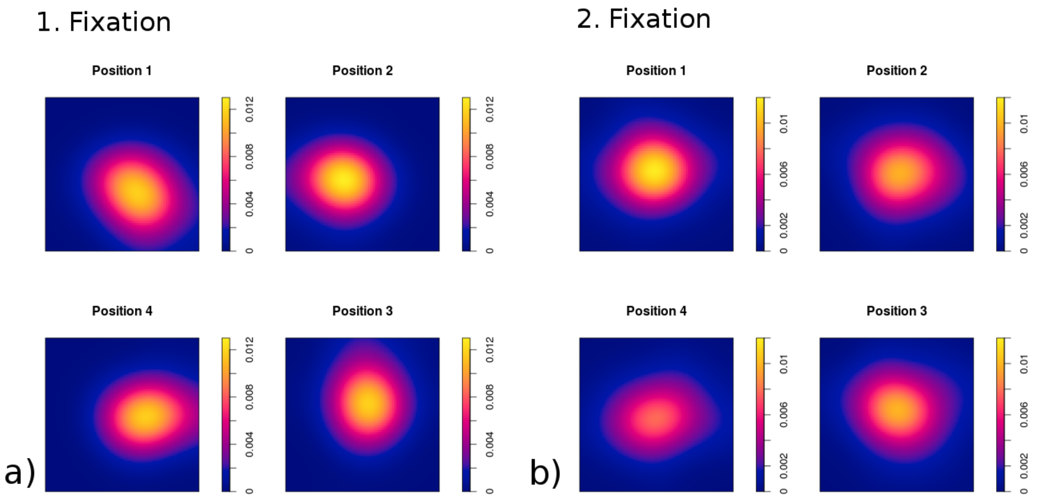

Figure 16: Visible effect of the quadrant design. In the first fixation the CFB is distorted in the direction of saccade origin. This effect disappears for the second fixation.

4.3 Preview vs. Flash

It is not be possible in the scope of this thesis to draw conclusions concerning the distinct effects of the onset type and the preview. Our sample size of 5 subjects is not large enough to reasonably conduct such tests. In a following larger scale study, the results should be analyzed using a within subjects ANOVA with the factors of preview and onset.

We were, however, able to find some relevant tendencies in the visualizations. Fig. 17 shows the results of the 2x2 design.

The weakest CFB was found for the *control preview no flash* condition, followed by the standard *preview/no flash* condition. This reproduces the effect found by Rothkegel et al..

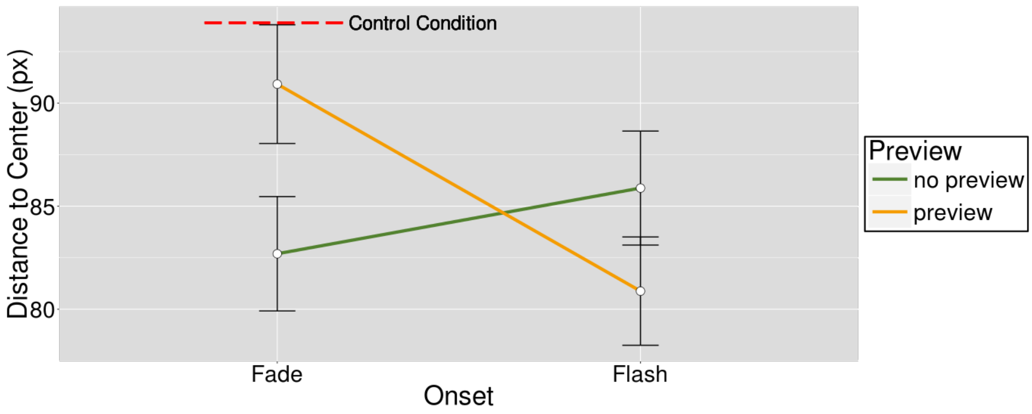

Figure 17: The effects of onset and preview on the central fixation bias. A smaller average distance to the center represents a stronger CFB.

5 Discussion

In this thesis we proposed an experiment to dissociate between two possible explanations for the central fixation bias. One examined possibility is that the first fixations are made in the center of an image because it is the optimal location to extract gist information about the image. In this case, if the viewer had such information prior to freely exploring the image, the central fixation bias should be reduced. The second explanation is that the luminance change of the screen that is associated with the sudden presentation of images attracts the eyes. This would allow a reduction in the central fixation bias by presenting the images without sudden onset.

The proposed experiment implemented the presence or absence of gist extraction by allowing or disallowing peripheral previews, using a 2x2 grid to present images. Images were either peripherally present, allowing viewers to extract information before viewing, or appeared only when they were to be viewed directly, without preview.

The second experimental factor was the type of onset. Sudden luminance change was implemented as the usual all-at-once presentation, while its absence was operationalized as a fade-in effect. We also included a fifth condition to control for the unusual use of the fade mechanic. It did not include a fade or flash effect, but instead presented the entire stimulus continuously from the beginning of the trial.

The main objective of this thesis is to assess whether the described approach is suitable for researching the central fixation bias. In this section we will discuss the gathered results and answer the research questions defined in the Introduction (sec. 1).

5.1 Preview Implementation

The data concerning viewing order compliance (sec. 4.1) indicates that participants were able to adhere to the rules of the quadrant design. Subjects showed a consistent amount of saccades over the positions; there were no preferred positions and only relatively few deviations from the prescribed order.

The only unexpected result in this regard was the particularly large amount of saccades going back to previous images, which was evoked by the control condition in some participants. This condition differed from all the others in that it did not include a fade or flash effect. It was also the last condition for all participants. The large amount of backwards jumps could be due to this difference, especially considering that participants spent the four previous conditions getting used to a consistent routine.

We found a central fixation bias, which, as expected, was strongest in the first two fixations. We also found that when including very short saccades in the analysis, the second fixation was closer to the center than the first, but that this effect disappeared when disregarding fixations shorter than 150 ms. A possible explanation for this phenomenon is that the length of the saccade needed to jump from one image quadrant to the next exceeded the usual saccade length. Subjects may, therefore, have not reached the intended fixation position (in the center of the image) but landed slightly off, causing them to compensate by quickly making a corrective saccade. This idea is further supported by the observed landing effects. The saccade origin determined in which direction the central fixation bias of the first fixation in the next image was distorted.

Most of these effects of the quadrant design are expected and are not likely to be detrimental to the quality of the data gathered by the experiment when performed on a larger scale. Taking this into consideration, we can answer RQ1 in the affirmative. The quadrant design appears adequate to test the effect of the gist extraction on the central fixation bias.

5.2 Onset Implementation

To understand the effects of the fade mechanic, we must examine how the individual conditions affected the viewing behavior.

Out of the four main experimental conditions, participants showed the weakest central fixation tendency in the *preview no flash* condition. This condition was also the closest to the *control preview no flash* condition. The fact that these two conditions evoke similarly strong central fixation biases indicates that they are two implementations of the same *no flash* onset type and their mechanisms can be equated.

We can therefore tentatively endorse the fade effect as a suitable method to implement presentation of images without sudden luminance change (RQ2).

However, it's use in the *preview flash* condition of this experiment may need to be reviewed, as discussed in the next section.

5.3 Causes for the Central Fixation Bias

The collected data suggests a confirmation of the finding that information about the image (preview) and the absence of a sudden onset (flash) yield results with the least strong CFB, as also found by Rothkegel et al. The *preview no flash* condition caused participants to exhibit the weakest CFB.

All conditions, especially the three conditions with the strongest CFB, are too close together to claim significant differences at this point. If we regard the two *no preview* conditions as not significantly different from one another, the results indicate that image onset has an influence on the central fixation bias only when the participants are shown a preview. This is the case because the *no preview flash* condition, which we expected to have the strongest CFB, only displayed the third strongest bias. This could be due to a weakness in the experimental design. In the *preview flash* condition one could argue that subjects experience two sudden luminance changes (or flashes): one when the previously present image in the periphery disappears, and one when it reappears. This is a possible reason for the strong bias in the *preview flash* condition.

In most respects the experiment was conducted successfully. Concerning RQ3, we are hopeful that this experimental design will help us gain insight into the causes of the central fixation bias. However some limitations of the design ought to be improved in the future.

5.4 Limitations and Future Work

Firstly, this experiment depended largely on subjects complying with the given instructions. As described above, some subjects tended to deviate from the prescribed order, which may have compromised the results. A possible solution to this issue is to include a fixation check, which allows the stimulus to be seen only when the subject's eyes are in the intended quadrant.

In future implementations of this design we will randomize the blocked conditions. Having them always appear in the same order may have introduced artifacts into the results.

Another point to consider is whether the disappearance and reappearance of the stimulus in the *preview flash* condition can somehow be replaced. An implementation of this condition that avoids the double flash effect may clear up some uncertainties about the interpretation of results.

The use of the fade effect as an antagonist to the flash seems to have been successful. It would be interesting to develop an experimental design closer to that of the original experiment by Rothkegel et al., but controlling for the sudden luminance change by using a fade.

6 Conclusion

The central fixation bias is an effect often found in scene viewing experiments. Subjects tend to make their first fixations centrally in images. In this thesis we described a method for distinguishing two possible causes for the central fixation bias in scene viewing: sudden onset luminance change and gist extraction. The experiment consisted of a 2x2 grid of images shown with or without a preview, and with or without sudden onset. We conducted a pilot study with 5 subjects to assess the validity of this method. The implementation of the experiment was successful. Subjects were able to follow the quadrant design and deviated from the prescribed viewing order relatively rarely. The fade mechanic also seems to have been a legitimate implementation of presenting images without sudden luminance changes. We were able to tentatively reproduce the effects found by Rothkegel et al., which suggest that a preview and no sudden onset image presentation reduces the central fixation bias. The results of this thesis are encouraging for conducting this experiment with a larger subject base.

References

Arons, B. (1992). A Review of The Cocktail Party Effect. *Journal of the American Voice I/O Society*, *12*(7), 35—-50.

Bacon-Macé, N., Macé, M. J. M., Fabre-Thorpe, M., & Thorpe, S. J. (2005). The time course of visual processing: Backward masking and natural scene categorisation. *Vision Research*, *45*(11), 1459–1469. doi: 10.1016/j.visres.2005.01.004

Bahill, A., Adler, D., & Stark, L. (1975). Most naturally occurring human saccades have magnitudes of 15 degrees or less. *Investigative ophthalmology*, *14*(6), 468–469.

Bear, M. F., Conners, B. W., & Paradiso, M. A. (2007). *Neuroscience: Exploring the Brain* (2nd ed.).

Brainard, D. H. (1997). *The psychophysics toolbox.* Retrieved from `http://psychtoolbox.org/`

Broadbent, D. E. (1954). The role of auditory localization in attention and memory span. *Journal of Experimental Psychology*, *47*(3), 191–196. doi: 10.1037/h0054182

Buswell, G. T. (1935). *How People Look at Pictures.* (Vol. 33) (No. 2). Chicago: Univ. Chicago Press.

Cherry, E. C. (1953). Some Experiments on the Recognition of Speech, with One and with Two Ears. *The Journal of the Acoustical Society of America*, *25*(5), 975. doi: 10.1121/1.1907229

Engbert, R., & Kliegl, R. (2003, apr). Microsaccades uncover the orientation of covert attention. *Vision Research*, *43*(9), 1035–1045. doi: 10.1016/S0042-6989(03)00084-1

Foulsham, T., & Kingstone, A. (2012). Modelling the influence of central and peripheral information on saccade biases in gaze-contingent scene viewing. *Visual Cognition*, *20*(June 2015), 546–579. doi: 10.1080/13506285.2012.680934

Henderson, J. M., Weeks, Phillip A., J., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and*

Performance, *25*(1), 210–228. doi: 10.1037/0096-1523.25.1.210

Hoffman, J. E. (1998). *Visual attention and eye movements.*

Irwin, D. E., Colcombe, A. M., Kramer, A. F., & Hahn, S. (2000). Attentional and oculomotor capture by onset, luminance and color singletons. In *Vision research* (Vol. 40, pp. 1443–1458). doi: 10.1016/S0042-6989(00)00030-4

Itti, L., Koch, C., & Niebur, E. (1998). A Model of Saliency-Based Visual Attention for Rapic Scene Analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*(11), 1254–1259. doi: 10.1109/34.730558

Judd, T., Ehinger, K., Durand, F., & Torralba, A. (2009, sep). Learning to predict where humans look. In *2009 ieee 12th international conference on computer vision* (pp. 2106–2113). IEEE. doi: 10.1109/ICCV.2009.5459462

Loftus, G. R., & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance*, *4*(4), 565–572. doi: 10.1037/0096-1523.4.4.565

Mackworth, N. H., & Morandi, A. J. (1967, nov). The gaze selects informative details within pictures. *Perception & Psychophysics*, *2*(11), 547–552. doi: 10.3758/BF03210264

Mannan, S. K., Wooding, D. S., & Ruddock, K. H. (1996, jan). The relationship between the locations of spatial features and those of fixations made during visual examination of briefly presented images. *Spatial Vision*, *10*(3), 165–188. doi: 10.1163/156856896X00123

MATLAB R2015a. (2015). Natick, Massachusetts: The MathWorks Inc.

Parkhurst, D., & Niebur, E. (2003, jun). Scene content selected by active vision. *Spatial Vision*, *16*(2), 125–154. doi: 10.1163/15685680360511645

Posner, M. I. (1980, feb). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25. doi: 10.1080/00335558008248231

Reinagel, P., & Zador, A. M. (1999, jan). Natural scene statistics at the centre of gaze. *Network: Computation in Neural Systems*, *10*(4), 341–350. doi: 10.1088/0954-898X_10_4_304

Rentschler, I., Hauske, G., Schill, K., Zetzsche, C., & Krieger, G. (2000, nov). Object and scene analysis by saccadic eye-movements: an investigation with higher-order statistics. *Spatial Vision*, *13*(2), 201–214. doi: 10.1163/156856800741216

Rothkegel, L. O. M., Trukenbrod, H. A., Schütt, H., Wichmann, F. A., & Engbert, R. (2016). Sudden image onset generates the central fixation bias in scene viewing. *Unpublished manuscript*.

Salthouse, T. A., & Ellis, C. L. (1980). Determinants of eye-fixation duration. *The American journal of psychology*, *93*(2), 207–234. doi: 10.2307/1422228

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: sustained inattentional blindness for dynamic events. *Perception*, *28*(9), 1059–1074. doi: 10.1068/p2952

Smith, T. J., & Henderson, J. M. (2009, aug). Facilitation of return during scene viewing. *Visual Cognition*, *17*(6-7), 1083–1108. doi: 10.1080/13506280802678557

Tatler, B. W. (2007). The central fixation bias in scene viewing: selecting an optimal viewing position independently of motor biases and image feature distributions. *Journal of vision*, *7*(14), 4.1–17. doi: 10.1167/7.14.4

Tatler, B. W., Baddeley, R. J., & Vincent, B. T. (2006, jun). The long and the short of it: Spatial statistics at fixation vary with saccade amplitude and task. *Vision Research*, *46*(12), 1857–1862. doi: 10.1016/j.visres.2005.12.005

Theeuwes, J. (1995, jul). Abrupt luminance change pops out; abrupt color change does not. *Perception & Psychophysics*, *57*(5), 637–644. doi: 10.3758/BF03213269

Theeuwes, J., Kramer, A. F., Hahn, S., & Irwin, D. E. (1998, sep). Our Eyes do Not Always Go Where we Want Them to Go: Capture of the Eyes by New Objects. *Psychological Science*, *9*(5), 379–385. doi: 10.1111/1467-9280.00071

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes:

The role of global features in object search. *Psychological Review*, *113*(4), 766–786. doi: 10.1037/0033-295X.113.4.766

Vitu, F., Kapoula, Z., Lancelin, D., & Lavigne, F. (2004, feb). Eye movements in reading isolated words: evidence for strong biases towards the center of the screen. *Vision Research*, *44*(3), 321–338. doi: 10.1016/j.visres.2003.06.002

Võ, M. L.-H., & Henderson, J. M. (2010). The time course of initial scene processing for eye movement guidance in natural scene search. *Journal of vision*, *10*(3), 14.1–13. doi: 10.1167/10.3.14

Wood, N., & Cowan, N. (1995). *The cocktail party phenomenon revisited: how frequent are attention shifts to one's name in an irrelevant auditory channel?* (Vol. 21) (No. 1). doi: 10.1037/0278-7393.21.1.255

Yantis, S., & Jonides, J. (1996). Attentional capture by abrupt onsets: New perceptual objects or visual masking? *Journal of Experimental Psychology: Human Perception and Performance*, *22*(6), 1505–1513. doi: 10.1037/0096-1523.22.6.1505

Yarbus, A. L. (1967). Eye movements and vision. *Neuropsychologia*, *6*(4), 222.

Eigenständigkeitserklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbständig verfasst, sowie keine anderen Quellen und Hilfsmittel als die angegebenen benutzt habe.

Potsdam, 21. Juni 2016

Lisa Schwetlick