Visualization Research Center (VISUS)

University of Stuttgart
Allmandring 19
D–70569 Stuttgart

Masterarbeit

Influence of Distractors on Visual Attention Techniques

Nina Pascale Doerr

Course of Study: Informatik

Examiner: Prof. Dr. Michael Sedlmair

Supervisor: Katrin Angerbauer, M.Sc.

Commenced: December 15, 2021

Completed: June 15, 2022

Acknowledgements

I would like to thank my supervisors Katrin and Michael for the support throughout the thesis. Thank you for providing me feedback and motivating me when I got stuck. Although the last six months were demanding, I really enjoyed working on this. I am also thankful for the support, the coffee and the lunch breaks from all the other people of the AR/VR research group and VISUS. In particular, I would like to say thanks to those who provided me further inspirations in the pilot study. In that context, I am especially grateful for the possibility to work at VISUS for the last three years and hence, a big thank you to Michael B. and Guido, it was a pleasure. Thanks to my family and friends for all the support and also (as for all the other participants) for taking part in my study. Also a big thanks to my boyfriend for repeatedly testing the application, proof reading, bringing me food and giving me advices when I needed them. Finally, I would also like to thank myself that we made it happen!

Abstract

Many tasks in our daily routine require that we focus on specific parts while hiding others. Particularly for visualisation and in graphics, it is crucial that data representations are correctly perceived and the viewer can focus on key patterns and structures. In order to ensure this, there are several methods to direct the attention efficiently and to highlight significant elements, so-called visual attention techniques. Graphical systems like tablets or other 2D displays have a more restricted display space and it is common to apply efficient and computationally generated methods such as colour or shape variations to highlight objects of interest. With the application of augmented and virtual reality, the presentation space increases to 360° around the viewer and beyond. Therefore, we need techniques that direct the user to areas outside the field of view without disturbing the perception of a realistic world. Previous research mostly studied these types of techniques in less immersive, static scenes without any additional, potentially distracting elements (distractors). Thus, this paper aims to extend previous research and compares different visual attention techniques in terms of varying degrees of visual distraction. It investigates whether the influence of static (non-moving) or dynamic (moving) objects has an effect on perception and attention guidance. Results of this work show that distractors influence the perception of visual attention techniques to some extent. Furthermore, it is found that some techniques are more suitable than others in certain contexts. However, it still unclear which factors of distractors play a role in influencing the attentional guidance of the techniques.

Kurzfassung

Viele Aufgaben in unserer täglichen Routine erfordern, dass wir uns auf bestimmte Bereiche konzentrieren und andere ausblenden. Vor allem bei der Visualisierung und in der Grafik ist es entscheidend, dass Datendarstellungen richtig wahrgenommen werden und der Betrachter sich auf wichtige Muster und Strukturen konzentrieren kann. Um dies zu gewährleisten, gibt es verschiedene Methoden, um die Aufmerksamkeit effizient zu lenken und wichtige Elemente hervorzuheben, sogenannte visuelle Aufmerksamkeitstechniken. Grafische Systeme wie Tablets oder andere 2D-Displays haben einen begrenzten Darstellungsbereich und es ist üblich, effiziente und computergenerierte Methoden wie Farb- oder Formvariationen anzuwenden, um interessante Objekte hervorzuheben. Mit der Anwendung von AR oder VR vergrößert sich der Darstellungsraum auf 360° um den Betrachter herum und darüber hinaus. Daher benötigt es Techniken, die den Betrachter auf Bereiche außerhalb des Sichtfeldes lenken, ohne die Wahrnehmung einer realistischen Welt zu stören. Bisherige Forschungen haben diese Art von Techniken meist in weniger immersiven, statischen Szenen ohne zusätzliche, potenziell ablenkende Elemente (Ablenkungsfaktoren) untersucht. Das Ziel dieser Arbeit ist es daher, die bisherige Forschung zu erweitern und verschiedene visuelle Aufmerksamkeitstechniken in Bezug auf unterschiedliche Grade der visuellen Ablenkung zu vergleichen. Es wird untersucht, ob der Einfluss von statischen (sich nicht bewegenden) oder dynamischen (sich bewegenden) Objekten eine Auswirkung auf die Wahrnehmung und die Aufmerksamkeitslenkung hat. Die Ergebnisse dieser Arbeit zeigen, dass Ablenkungsfaktoren die Wahrnehmung von visuellen Aufmerksamkeitstechniken in gewissem Maße beeinflussen. Außerdem ist feststellbar, dass einige Techniken in bestimmten Kontexten besser geeignet sind als andere. Es ist jedoch immer noch unklar, welche ablenkenden Faktoren eine Rolle bei der Beeinflussung der Aufmerksamkeitslenkung der Techniken spielen.

Contents

List of Figures

List of Tables

1 Introduction

While driving in your car to meet some friends, several other cars come from different directions and pedestrians cross the street. Perhaps it is raining and some pigeons pass your wind shield while the radio plays some new songs and your friends send notifications to your mobile phone. Although this can be quite demanding, in the end you will reach your destination. This is possible because humans can focus and shift their attention to certain areas of interest (e.g., pedestrians crossing the street) while ignoring others (e.g., phone notifications) [Car18; WH04].

Research on visual perception and attention is a major scientific area in several domains like neurology and psychology [Inc22; LF98; Max22; McL18; MK09]. Nevertheless, there is also a growing interest in other research areas including visualisation and graphics [CM84; HE12]. Since these domains study ways to visually represent large datasets, it is necessary to consider how humans perceive certain presentations and how the attention can be directed to important structures and patterns within the data. Particularly in the data representation on 2D displays, there exist several approaches and techniques which can efficiently direct the attention by using preattentive visual features or motion [WH04; WLB+14].

With the development of augmented reality (AR) and virtual reality (VR), new opportunities but also challenges regarding the data presentation and visualisation arose. While there is now the opportunity to consider and visualise the data in a 3D context, it is possible that the viewer may feel overwhelmed or misses important patterns [GAM18; GSEM17]. Besides information and scientific visualisation, VR also brought new possibilities for entertainment by using the 360° space for presenting movies (so-called Cinematic VR) [RBH19]. However, all of these application areas require the viewer to focus on the relevant content while potentially ignoring the rest. Hence, it is even more relevant to efficiently direct the attention of the viewer. Since Cinematic VR aims to preserve immersion and a sense of presence within the virtual environment, it requires not only an efficient approach but also one that does fit within the presented environment [RH18].

While there exist several work regarding visual attention guidance in 2D environments, there is only few considering visual attention techniques in AR or VR. However, most of the previous work only considered whether their technique can direct the attention more accurately than others. Hence, the virtual environment consisted only of the technique, the target and partly, an immersive background [GSEM17; GTA+19]. Since real world surroundings mostly also include several other entities within the scene and especially in movies it can be quite busy, it is important to investigate potential techniques within such environments.

As there is barely work considering such environments, this work aims to compare and evaluate the influence of distracting scene objects on visual attention techniques. For this, we present in this work an implementation of our application using several additional scene objects (distractors) as well as a user study which studies this. Therefore, we selected the techniques Deadeye [KK19], HiveFive [LSGB20] and Circle (a combination of two preattentive features, [WH04]) which were previously shown to efficiently guide the attention. In addition, we used three different levels of distractors consisting of other non-moving and moving objects to evaluate if there are differences between scene objects. To study the influence, we considered the performance by measuring the completion time and the wrong clicked objects for each level of distractor of all techniques. In addition, we gathered some data regarding the perceived workload and subjective perception on these techniques by using own questions and NASA TLX. According to our results, using additional objects has an effect on the perception of visual attention techniques. While some techniques perform rather constant regarding various distractors, others are influenced by some specific distractor objects. However, more research is required to find out which type of distractor effects a technique and its reasons. It is also advisable to choose certain techniques wisely by considering its surrounding beforehand.

This work is structured as follows. Firstly, Chapter 2 summarises a general overview of visual perception and visual attention and examines previous work. For the latter, it is distinguished between representations on 2D display and in AR and VR as well as research focusing on distractors. Afterwards, the implementation of our virtual study application and its components is presented in Chapter 3. Subsequently, Chapter 4 provides an overview of the structure and conduction of the user study. This is followed by the presentation and discussion of our user study results in Chapter 5. Finally, Chapter 6 summarises the most important findings and limitations of our work and gives an outlook on future work.

2 Background and Related Work

As there is several research on the human brain and vision, it is important to generally describe this in advance. Hence, this chapter first introduces the domain of visual perception and visual attention. Beyond that, we also review some previous work on attention techniques for 2D and immersive environments (AR and VR). Given that this work aims to compare such techniques regarding distractor interference, the last section also summarises previous research that examined the influence of distractors on attention and more precisely, attention guidance.

2.1 Visual Perception

As a human, we have several sensory organs (e.g., eyes or nose) and each is part of a sensory system [McL18]. With these sensory organs, we can detect information from our environment. The associated sensory system then receives these sensory information and transmits them to our brain. Our brain then converts them in some ways so that we perceive certain sights or smells. The process of understanding and interpreting what our eyes see is called visual perception [Gre20; Inc22]. Due to visual perception, it is possible to perform several tasks. For example, when we are driving a car, we have to perceive the presence of other cars, their distance and speed. While there is several research of various psychologists and neurologists of the process connecting the senses and our perception of the surrounding, visual perception is also essential for visualisation and graphics [HE12; Max22]. By understanding the process of visual perception, it is possible to optimize the information representation qualitatively and quantitatively.

The work by Cleveland and McGill [CM84] is among the first studies on visual perception in terms of graphical methods for data presentation and analysis. Even though William Playfair (according to Cleveland and McGill [CM84]) already used graphs to visualize data in 1786, during the time of the work of Cleveland and McGill [CM84], the design of graphs were a matter of common sense without any scientific foundation. Hence, they aimed to provide such a foundation using graphical perception, the visual extraction of information represented in graphs. Within their work they identify several elementary perceptual tasks that are used to extract quantitative information in graphs and reorder them based on the accuracy of humans. These elementary perceptual tasks are depicted in Figure 2.1. For example, judging the position, length or area (e.g., in bar charts or scatter-plots) to identify the relative magnitude of a certain data point. However, these tasks can also be related to each other (e.g., the judgment of angle and direction). In addition, while they also include shading, they exclude the colour hue or texture as these can not be ordered unambiguous and hence, rather encode categories. Their order is based on the accuracy of extracting quantitative information such that a higher accuracy of one task means that the perceived value is closer to the actual value than for another task. They also expect that a representation with a more accurate task also increases the probability of perceiving hidden patterns or structures. Based on results of psychophysics and their own reasoning, they claim that the position is perceived more accurate

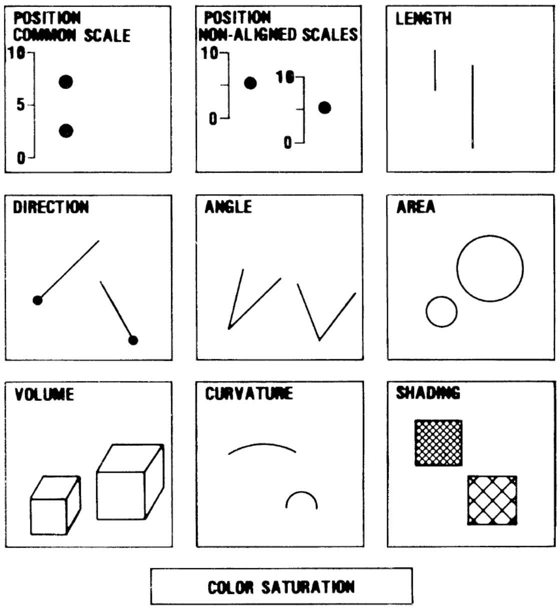


Figure 2.1: Elementary perceptual tasks (ordered from left to right, top to bottom) [CM84].

than length, direction or the angle. Subsequently, these are perceived more accurate than area, volume or shading. In addition, they perform experiments to measure the actual accuracy to validate their theoretical findings. Their results show that position judgments are more accurate than judgements of length or area and also indicate a constant underestimation of values for length and angle judgments. Their insights intended to aid re-designs or designs of graphical forms so that the data can be perceived more accurately. For example, they suggest to replace certain types of graphs by others (pie charts by bar charts) [CM84]. While the previous work focused more on a traditional 2D application, there is also work that investigates these findings using immersive virtual and augmented reality displays [WSA20].

2.2 Visual Attention

An important component of visual perception is the visual attention [Car18; DBD07]. Despite the fact that we are continuously exposed to an overwhelming quantity of visual information during our waking hours, we can effortlessly understand our visual environment and identify and localize relevant information. Since our visual system has a finite processing capacity, it cannot process all present information at once but uses some kind of selection (e.g., it limits the processing to a subset of perceivable objects or accessible memories) [LF98; MK09]. This selective processing results from using attention to focus on certain regions and hence, prioritises these information [Car18]. Thus, we can define visual attention as a collection of cognitive processes to select relevant information and to filter out irrelevant information of a visual scene [MK09]. This means that we are only aware of the parts of a scene that we pay attention to or that attract our attention [DBD07]. Carrasco [Car18] distinguishes between overt attention (by moving eyes to a certain area) and covert attention (by paying attention to a peripheral region but not actually directing eyes to it).

For instance, when driving in a car, for overt attention you would directly look behind you when overtaking someone whereas for covert attention you attend to traffic lights while still considering the road [MK09].

Besides this, Healey and Enns [HE12] review and summarise several research findings on visual attention and point out its relevance to information and scientific visualisation which are also important for our work. For visualisations and our work, it is necessary to understand how our visual system analyses images. In general, we only have a detailed vision within a small part of the visual field. To gather further detailed information in other regions, our eyes unconsciously alternate between so-called fixations (to acquire new details) and saccades (flickering of the eyes to a new location within a period of blindness).

According to Healey and Enns [HE12] several previous research of visual attention considered the consequences of the selection of specific objects or locations. Besides several theories on visual attention, one important finding for visualisation research was the identification of preattentive features. These are a limited set of visual features and called preattentive as they can be detected within a single fixation before the first saccadic movement is triggered (< 200-250 ms). An example task is finding a red circle within many blue circles which can be seen in Figure 2.2. Due to its unique visual property (red instead of blue colour), it pops out of the display and is easy to find regardless of the amount of distractor objects (blue circles). However, if a target has more than one visual property (so-called conjunction of features), they cannot be detected preattentively and the observer has to perform serial searches. For example, when searching a red circular target within distractors that have either one of these features. Besides the identification of preattentive processing, Healey and Enns [HE12] summarise several theories which tried to explain how the visual system performs preattentive processing. Within this work, we only present the feature integration theory because one of our visual attention techniques (Deadeye [KK19], see Section 3.3) is based on this.

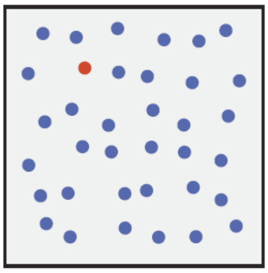


Figure 2.2: Target detection: red circle target among blue circle distractors [HE12].

As one of the first visual attention researchers, Treisman (according to Healey and Enns [HE12]) analysed the preattentive processing and focused on the classification of these visual features. For this, they conducted experiments with target and boundary detection tasks to obtain a set of visual features that are perceived preattentively. Within their experiments they measured the performance of the participants either by response time or accuracy. By using the response time, a visual feature is preattentive when the completion time of the task is constantly below a certain threshold regardless of the amount of other objects within the scene. When using accuracy, a target is either present or absent for a certain duration. After the time expired, a participant states the current state of the target and hence, the visual feature of the target is preattentive if their predictions are accurate.

In addition, Treisman (according to Healey and Enns [HE12]) gave a potential explanation on how the visual system handles the preattentive processing by describing the human low-level vision with a model consisting of several feature maps and a master map of locations. With this model, it is possible to explain preattentive processing as well as the conjunction search. They also extended their feature integration theory to a spectrum consisting of parallel or serial feature detection and gave explanations to situations in which a conjunction search is also perceived preattentive. Both extensions depend on a significant target-nontarget difference.

Besides this, two other concepts are relevant in terms of preattentive visual features. The first is called ensemble encoding of visual features. It indicates that humans are able to summarise the distribution of visual features seen in an image but cannot state if an object had a specific visual feature like a certain size. This is also the case for colour, orientation and other higher-level qualities like emotions. The other concept is a feature hierarchy suggesting that some visual features are more salient in some tasks (e.g., colour over shape for boundary detection tasks). This can be very beneficial in terms of visualising data without masking any information.

Besides the representation of the data (e.g., using preattentive features), it is known that also our goals and tasks guide saccades to new locations. Hence, our visual attention is influenced by more than just the visualisation when selecting the next image region for further processing. There also exist several research findings of visual attention which investigated the amount of information that humans can recall after looking at a certain scene. By understanding the influence of what humans think or remember while considering a certain visual representation, it is possible to design visualisations that direct and support the viewer to locate and focus on the most important data. For example, some researchers found a phenomena named postattentive amnesia suggesting that there is no improvement in search time if a person can look at a scene prior to searching for a certain target. In addition, there is some research indicating that memory can in some way be helpful for search time. For example, by using subtle regularities the attention can be directed to a certain target (so-called contextual cuing). This means the unconscious knowledge of the global context (e.g., spatial relations of target and distractors) influences the perception. Furthermore, it is shown that humans have an unconscious tendency to rather search targets in unknown than in previously attended locations. However, this tendency significantly differs to known, conscious strategic search patterns. Further research, summarised by Healey and Enns [HE12], demonstrates that a search interruption can be faster resumed than starting a new one. Another phenomena is change blindness, indicating that humans are blind to significant changes which occurred during an interruption of a viewed scene. Common examples are pairs of images, with one image differing in certain aspects. While there exist several explanations to this phenomena, the origin as well as the reasons why some changes are detected and other remain unnoticed is still a research topic. Hence, it might be useful to be aware of the attention of the viewer to avoid change blindness.

Besides change blindness, there is also a phenomenon called inattentional blindness. It indicates that people tend to fail in perceiving salient objects even when being directly within the field of view. For example in one experiment, participants had to watch a video and their task was to count bounces [HE12]. Within the video sequence, an unexpected event (a woman with an umbrella walked through the scene) occurred, however several participants did not notice the woman. Some research attempted to guarantee that viewers do not miss relevant visualisation changes as unawareness could have some potential consequences. With an increase of the screen size the probability of missing relevant changes increases. However, it is possible for small and larger screens that viewers notice changes when being in the focus of attention while remaining unnoticed if the scene impression stays the same. There are also some approaches that attempted to benefit from perceptual blindness, for example to reduce rendering costs. There exists also research considering the availability of attention over time (named attentional blink). It has been found that a subsequent target is not perceptible when being presented within 100 to 500 ms after the original target. Therefore, participants failed to report the presence or absence of the subsequent target although being accurate after this interval as well as immediately after the first target. To summarise this in terms of visualisations or other visual data representations, it is important to know that some changes can be overlooked if a viewer does not directly focus their attention on this data.

According to Healey and Enns [HE12], there are more ways to benefit from visual attention in terms of visualisation. Firstly, it is possible to track the attention or forecast the gaze allowing to design visual representations based on the expected amount of attention. For example, by only rendering all details which a viewer most likely focusses on. This is especially helpful for photorealistic rendering to spend less time on rendering. More importantly for this work, there exist several visualisation techniques that can direct the attention of the viewer. This is beneficial for visualisation to highlight key data points while other data values can still be considered. In addition, using the right visual features the attention of a viewer can also be guided to essential areas of interest within the data. For example, there are methods for recolouring images such that colour-blind viewers can perceive differences in colour and hence, can consider such visualisations. To sum this up, research on the proper usage of attention still continues leading to significant implications for visualisations. Since the goal remains to create salient and memorable graphical representations which guide the attention to key data points, it also remains crucial to understand the visual system [HE12].

As mentioned earlier, it is possible to guide the attention with some visualisation techniques to highlight important data points. According to Healey and Enns [HE12], guiding the attention in visualisations to its key data points can enhance engagements and insights of a viewer. However, it is necessary to find approaches that do not confuse the viewer. Several researchers presented already efficient methods or techniques, so-called visual attention techniques in context of visualisation or graphical representations for 2D but also for AR and VR.

2.3 Visual Attention Techniques in 2D

There are several influences which guide the gaze of a person if they are looking at a traditional static 2D image [BMSG09]. Besides subjective goals or tasks, the depicted content (e.g., faces, informative parts) can affect the eye movement pattern as well. In addition, viewers can find targets faster and easier when being embedded within a familiar scene (e.g., locating a bread within a kitchen) [HE12]. Hence, it is reasonable to represent data within their correct semantic context such that the viewer's attention is directly guided to the most probable locations. However, while this is possible for some general representation, this is more difficult for artificial or scientific visualisations. These representations normally do not have a real semantic context. In general, this is beneficial for augmented or virtual reality applications. Also image properties like high local density or contrast direct the gaze [BMSG09]. Digital representations of images offer further possibilities to actively guide the attention to certain regions. This is especially relevant for cluttered visualisations or other graphical representations to actively highlight relevant content to the viewer so that they do not get lost in the data.

As previously mentioned, preattentive features can be detected within 200 to 250 ms regardless of the presented amount of distractors [HE12]. In addition, it is possible for some of these visual features to guide the attention of a viewer to a certain target [WU18]. Additionally, a combination of more than one feature can direct the attention efficiently, for example, using a target with shape and colour variations. Wolfe and Horowitz [WH04] suggest that not all visual features can guide the attention equally well or at all. Therefore, they provide an ordered list of visual features ranging from undoubted attributes to probable non-attributes. For the former, undoubted means that there is several evidence for potential attention guidance, while the latter (probable non-attributes) says that results indicate that these features do not guide at all. According to the list of Wolfe and Horowitz [WH04] colour, motion but also stereoscopic depth can guide the attention. Since our selected techniques include one to multiple attributes, we assume that our selected techniques are consistent under distractors.

Apart from preattentive visual features, there are also other visual approaches to guide the attention to areas of interest. Many previous approaches applied common naturally perceived methods for highlighting, for example, by simulating a depth-of-field effect [BMSG09]. This effect leads to target areas appearing in focus while less important areas appear to be out of focus. Such an effect can be achieved by applying a sharpening filter and a blurring filter which both are available in image editing packages. Another benefit besides the availability of such a technique is that it is also applicable for 3D scenes. Additionally, Healey and Enns [HE12] refer to a pen-and-ink sketch approach that directs the attention by using a high level of detail for frequently varying data values and only some strokes for rather constant data values. Another study compared the gaze patterns for the rendering types of photographs by using an image with full detail rendering, a non-photorealistic rendering and a combination of both rendering styles with only full detail rendering for target locations. Due to the results showing that participants focus on target locations even for the third rendering style, it is possible that reasonable rendering abstractions can guide attention. This is also found for further aesthetic variances depending on data values. According to Bailey et al. [BMSG09], a similar approach also applied stylized rendering for attention guidance in a 3D scene. However, such a visual modification can result in unintended misinterpretations. It might be unfavourable to use a whole visual dimension for highlighting. There exists research using stereoscopic depth in 2D or 3D graph visualisations to project certain data aspects closer to the viewer and hence, visual features like colour or size can be used for other representations [KK19]. However, it is difficult to perceive depth accurately.

As discussed earlier, Wolfe and Horowitz [WH04] also listed motion besides colour or orientation as a potential feature that can direct the attention. Due to their own technique, Waldner et al. [WLB+14] summarise several approaches that studied motion and flicker as potential attention guidance techniques. One of these techniques used a blinking target to demonstrate that it can be distinguished from moving distractors [POT08]. Since we also consider dynamic distractors in our study, we will later consider their distractors. According to Waldner et al. [WLB+14], some of their summarised studies showed that motion is an effective feature to guide attention, but there are some results indicating that participants find moving attractors annoying. Therefore, they present and discuss another attention guidance technique as a compromise of effectiveness and annoyance. Their so-called attractive flicker is based on luminance oscillations of the target consisting of two phases. While the first phase tries to draw the attention to the target points by short and intense flickering, the second phase intends to assist the viewer to monitor these targets. Hence, the short intense flickering is reduced to minimal luminance oscillations in the second phase and thus, does not annoy the viewer.

In contrast to the conscious and clearly perceivable attractive flicker, Bailey et al. [BMSG09] present a more subtle approach called Subtle Gaze Guidance (SGD) to guide the attention. For this, they exploit the fact that the foveal vision of humans has a very high acuity and the peripheral vision identifies certain stimuli quicker. Hence, our foveal vision tries to fixate what the peripheral vision recognized beforehand. Bailey et al. [BMSG09] tried to guide the attention by modulating target regions in the peripheral vision using either luminance (alternately interpolating black, white pixels) and a warm-cool (alternately interpolating red, blue pixels) modulations. They assumed that the peripheral vision recognizes these modulations and thus, our foveal vision tries to fixate them. However, they terminated these modulations before the foveal vision reached the target. For this, they monitored the eye movement by a real-time eye tracker. Before evaluating their approach, they determined a threshold such that the modulation is sufficiently intense but still subtle. They conducted a user study to determine whether their approach can direct the attention and identify potential implications on the perceived image quality. For this, they presented 40 images with manually selected regions using also rather uninteresting regions. Between each image, the participant had to fixate a black screen with a fixation cross. Besides the measured data, each participant had to give feedback on the perceived image quality. Their results showed that the modulation can attract the participants gaze and alter their viewing pattern especially with luminance modulations. In addition, participants rated the modulated image quality lower than images without modulations [BMSG09]. Besides purely visual solutions, there exists also further research using sound for attracting attention [MF95]. Since we did not apply such methods in our work, we will not consider this any further.

2.4 Visual Attention Techniques in Augmented and Virtual Reality

Immersive headsets became affordable consumer products with a wide range of applications and research from storytelling in entertainment to scientific data visualisation [GSEM17; LSGB20]. Recent advances in screen resolution and field of views also increased the perceived immersion and hence, it requires more than ever considerations on how to present information within these augmented or virtual environments. Particularly due to the missing frame and the enhanced area to illustrate the data (i.e., all around the user), users are likely to feel overwhelmed and overlook or miss relevant content [GAM18; GSEM17]. Thus, efficient techniques for visual attention guidance are even more relevant for these immersive environments. However, the majority of work still exists for static presentation forms (e.g., 2D image representations) and only limited research deals with visual attention in virtual 3D environments [HLB+16].

There exists research on applying common solutions which were shown to be very effective for the 2D domain (see Section 2.3), for example, by adding additional elements like arrows or adjusting visual properties (e.g., colour or brightness) of the target objects [LSGB20]. However, these features might be perceived unreal or unnatural and thus, reduce the intended immersion or sense of presence in a virtual environment. Since another major research area in VR is to increase the immersion and sense of presence to create ‘realistic’ virtual environments, such approaches to highlight important content might be unfavourable [NMH+16; RH18; SFR01]. As mentioned before, some research used a blurring filter so that the area of interest appears sharp and in focus [GTA+19]. Lange et al. [LSGB20] summarise several approaches including one that tried to adapt this technique for virtual environments. They state that while it can be perceived very quick, it is also perceived unnatural and has a degrading effect on immersion as well. Besides approaches that try to manually direct the gaze and attention of the user, there is also research that automatically changes the user's point of view [GTA+19; LSGB20]. For example, by rotating the virtual representation of the user towards the regions of interest even though the user can still turn their head arbitrarily. These approaches have a negative effect on the immersion and also need supplementary hardware.

Besides these rather additional cues to guide attention, there is research trying to guide the attention by objects that are sensed as a part of the scene. For example, an actor pointing towards the target area [GSEM17; GTA+19]. As mentioned before, motion is also an effective method for attention guidance [GSEM17; GTA+19]. A few approaches for 2D (see Section 2.3) but also for VR examined whether moving stimuli could attract attention of a viewer. Prior to presenting their own motion-based technique, Lange et al. [LSGB20] present an approach that integrated fireflies in a virtual environment which moved from the field of view of the user towards the area of interest. According to Lange et al. [LSGB20], this approach preserved the immersion and helped the user to follow the content without missing the content. They conclude that such techniques are beneficial to guide attention guidance in VR but also point out that, to their knowledge, there is no generic approach that can be applied in a variety of contexts. For this reason, they present their technique HiveFive which is one of the techniques we compare in this work. They investigate with their technique whether swarm motion can be applied to direct attention in general and also in peripheral regions. According to Lange et al. [LSGB20], their technique might be beneficial since swarms appear in various forms (e.g., fish or insects) in the real world and even formations of particles (e.g., dust or leaves) moving together may be perceived and used as a variation of swarms. Therefore, they claim one can apply a certain representation of a swarm depending on the chosen surrounding or environment to preserve immersion. In our case, we use the same surrounding and type of swarm as Lange et al. [LSGB20] for our study. As we compare this technique in our experiment, we later explain the technique and the corresponding evaluation of Lange et al. [LSGB20] in further detail in Section 3.3.

A more subtle technique to guide the attention is the extension of the SGD (see Section 2.3) for immersive environments by Grogorick et al. [GSEM17]. They try to adapt the technique so that it can also direct attention in cluttered virtual scenes without decreasing the perceived immersion. Due to the enlarged field of view and 360° presentation area they added to the traditional SGD the following two properties. Firstly, they added a dynamic positioning of their stimulus to integrate possible head movements to allow the target to be behind the viewer. Each time a target is outside of the field of view, the stimuli repeatedly moves to the screen edge in direction of the target. Additionally, Grogorick et al. [GSEM17] included some shape adjustments of the stimuli as the traditional circular shapes become thin ellipses in peripheral vision. Like Bailey et al. [BMSG09] for their traditional technique, Grogorick et al. [GSEM17] also conducted a perceptual study to approximate the stimulus size across different positions in the VR headset. For this, they took a static scene with targets within the field of view and a dynamic scene with possible targets behind the user so that the user had to move their head. In both cases the participants had to increase or decrease the size until the stimulus was barely noticeable. While we also use a dynamic scenario, our definition of dynamic differs from the definition of Grogorick et al. [GSEM17]. While we did compare the techniques using several moving distractor objects, they refer to dynamic as the user has to move (their heads) within the virtual environment. Additionally, our targets were not out of sight. Besides the perceptual study, Grogorick et al. [GSEM17] conducted a user study to evaluate their extended SGD using an immersive visual search task and the dynamic scene as virtual environment. Their participants had to find 20 spheres, the first ten targets without any guidance and the last ten targets using the extended SGD technique. Their results showed improving detection times with the technique suggesting that the SGD is applicable for virtual environments. Moreover, in a further study by Grogorick et al. [GAM18], they compared their extended SGD with other subtle gaze guidance techniques within a virtual environment. Besides the SGD, they used ColorDot (a technique using small red squares), SpatialBlur (using spatial blurring to only have sharp details in target regions), ZoomRect and ZoomCircle (a magnification of the target in a rectangular or respectively, circular shape form). As their results suggest, the SGD achieved better results for the target search than the other techniques.

As previously mentioned (see Section 2.3), there are also approaches that use stereoscopic depth to guide the visual attention. While this is rather difficult for 2D displays, stereo equipment as HMDs enable new possibilities for enhancing the target location without distorting objects [KK19]. In general, we use binocular disparity to extract depth information as our eyes perceive two images with an offset. Since immersive headsets render a scene for each eye, it is possible to exploit this by adding small differences to one of the two images [GTA+19]. Due to this differing information at a certain position, our brain perceives a conflict called binocular rivalry. This is beneficial as one may have an additional degree of freedom for visual encoding but also, as Krekhov and Krüger [KK19] could show, is perceived preattentively. There exist some approaches using binocular rivalry to guide the attention.

For instance, in another work by Grogorick et al. [GTA+19], they investigated whether such a visual stimulus can direct the attention. For their technique, they increased the brightness of the area of interest for one eye while decreasing it for the other eye (see Figure 2.3). In addition, they evaluated their technique using an HMD and a dome projection system. As in our user study, their targets were within the field of view. They located their stimulus in four different positions (on left, right, top or bottom of the fixation cross, see Figure 2.3) with varying sizes and intensities in various 360° panorama images and their participants had to identify the stimulus position. In addition, they used an eye-tracker to verify that the participants fixated the fixations cross during the whole study procedure. For their results, they evaluated all possible pairs of size and intensity to gather more insights for the overall perceived visibility and the recognition rate of the stimulus laid between 60 to 95 percent. In general, Grogorick et al. [GTA+19] found that in terms of visibility, the intensity of the stimulus is more dominant than the size. Due to the usage of various different panoramic scenes, they suggest that the stimulus can be used for several scenarios. However, they also suggest that a further evaluation might be reasonable to test the stimulus in dynamic scenes or in video sequences. Although we did not use this technique for our user study, we tried to investigate such a scenario within our work. In addition, Grogorick et al. [GTA+19] plan to further study the perceived discomfort and the effect of their technique on immersion. While Grogorick et al. [GTA+19] exploit binocular rivalry to highlight the target with different degrees of brightness, Krekhov and Krüger

Figure 2.3: Attention guidance based on binocular rivalry by decreasing the brightness on the left eye and increasing the brightness on the right eye [GTA+19].

[KK19] attempt this by showing the target object to one eye only. As for HiveFive by Lange et al. [LSGB20], since we use this technique within our user study, we will consider the technique and its evaluation in Section 3.3.

Since Cinematic Virtual Reality (CVR) is a huge application domain for immersive headsets, there are also several approaches that consider attention guidance [NMH+16; RBH19; RH18; Yu19]. In general, CVR is regarded as a novel type of VR storytelling [Yu19]. Especially for movies, it is important to not lose the attention of the viewer so that they do not miss important story parts and to trigger the desired emotional reaction. According to Yu [Yu19], some researchers tried using conventional movie approaches to guide the attention of the viewer. However, this is ineffective if the attention is scattered in 360°. Hence, Yu [Yu19] tried a new form of attention guidance. Instead of directly guiding the attention of the viewer to a certain location or point in VR, they took advantage of the scattered attention. Within their movie, the story progresses depending on the gaze of the viewer. There are also examples of such interactive 2D movies (e.g., Black Mirror: Bandersnatch). In contrast to this movie where the viewer interacts with a GUI to continue the story, in the movie of Yu [Yu19] the viewer does not know about the interaction using their gaze. This is according to them preserving the immersion.

Besides using VR, there is several research using attention guidance in augmented reality [HBB20; RP17; RP18; RP20]. As we solely compare techniques that have been investigated in VR and also conduct a user study in a virtual environment, we do not consider this any further. However, we believe that our user study might also be relevant for such applications. In general, there might be more distractors than in a full virtual environment, as in AR, both real and virtual objects occur simultaneously. Whereas you can still control distractors in VR in some ways, unpredictable distractors might occur more likely in AR. In general, we still encourage to further study the influence of distractors for attention guidance techniques in AR as there are several interesting application areas (e.g., navigation or way finding).

2.5 Distractors for Visual Attention Techniques

In the last two sections, we presented and discussed several techniques that have been shown to effectively guide the attention of users towards a certain target area using 2D displays and immersive headsets. As previously mentioned, techniques for virtual content usually try to preserve the immersion. While most of the presented techniques applied a rather realistic surrounding or background, they did not add any further virtual objects or effects which would normally appear in a similar real world environment. For example, the stereoscopic technique by Grogorick et al. [GTA+19] presented the target area different for each eye using realistic looking 360° panoramic images. However, they did not add any further visual objects which might appear in a similar real surrounding (e.g., one picture shows an underwater shot, a real environment would have animals such as fish or sharks). In order to achieve more of a real overall impression and also to enhance the sense of presence, integrating these types of entities may be beneficial. Since CVR is a huge application domain of VR and attention guidance, additional visual but also auditory objects might also appear and distract the user. While visual attention techniques aim to counteract these distractions, there is also the possibility that general effect and accuracy of such a technique decreases with additional objects and hence, does not guide the attention as effective as before.

Although several domains investigated the concept of attention and distractions, there is limited work considering distractors for attention guidance techniques [ODZK18]. Some research suggests that similar features of targets and distractors might influence the perception of targets among distractors and results in a higher demand. Hence, it is possible that similar features of a technique and distractors (e.g., similar moving patterns) influence the perception of the attention guidance by a certain technique. For example, Olk et al. [ODZK18] investigated a rather general paradigm of attention and distraction in an immersive virtual environment. Although they did not use an attention guidance technique, it is possible that based on their work we can obtain new insights which might be beneficial for further research on these techniques using distractors. They conducted two experiments, the first in an immersive virtual environment using a six-sided active CAVE system and shutter glasses and the second in order to compare the results of the first experiment. For this, they used the similar procedure and the same environment and stimulus using screenshots to present it on a 2D monitor. In both of their experiments, participants had to search for daily objects among distractors in a virtual kitchen and had to answer which target object out of two was present (e.g., a yoghurt, see Figure 2.4a or a soda can, see Figure 2.4b). The daily objects were real kitchen objects such as soda cans or apples and placed on the countertop of the virtual kitchen (see Figure 2.4). They also varied the discriminability of the target and distractor. In one condition, the target and distractor object had two distinct colours (high discriminability, see Figure 2.4b) and in the other condition both objects had the same colour (low discriminability, see Figure 2.4a). In general, Olk et al. [ODZK18] expected a higher response time for low discriminability. As expected, their results indicate a higher response time for a low target distractor discriminability and hence, they suggest that a similar colour of targets and distractors requires more search time. In the second experiment, they also tracked the eye movements of the participant as they expected more fixations until reaching the target for a low target distractor discriminability. The results of the second experiment for the target distractor discriminability were similar to the first experiment. As previously stated, they analysed the eye patterns of the participants and as expected, there were more fixations for a low than for a high target distractor discriminability. It is important to note, that Olk et al. [ODZK18] also measured and evaluated other factors of attention within these two user studies. However, since we are only interested in the influence of distractors on targets, we only report the findings for the effect of attention and distractors. Similar to their approach, some of our static distractors have a low target distractor discriminability (i.e., using trees with apples for target and distractor objects).

(a) Low discriminability between target (red yoghurt) and distractors (further red objects)

(b) High discriminability between target (red soda can) and distractors (further green objects)

Figure 2.4: Virtual countertop with targets among distractors [ODZK18].

By considering the work of preattentive features, an important requirement is that such a visual feature can be detected within 200 to 250 ms regardless of the presented amount of distractors [HE12]. Therefore, when investigating a potential preattentive feature it is evaluated using several distractors. In our work, we also compare techniques using static and dynamic objects.

McNamara et al. [MBG09] also examined the previously explained technique SGD ([BMSG09; GSEM17]) using static distractors in a 2D environment with two user studies. In the first study, they presented various images with 4 to 12 targets (using transparent spheres) which were uniformly distributed within the scene. In addition, they manually placed some targets at barely visible positions. They varied the image complexity, number of presented targets as well as the modulations of the target. They used three different modulation variations (no modulation at all, a subtle modulation as the original SGD and an obvious modulation by increasing the stimulus size of the original SGD). Their participants had to count the targets that were present for each image and verbally report them. In general, their results were quite consistent with a slightly higher accuracy when using modulations instead of using no modulations. Their results also indicated that modulations resulted in a higher percentage of returned counts. Furthermore, both modulation conditions yielded a higher percentage of correct answers. Moreover, McNamara et al. [MBG09] reported that participants did not notice the difference of the two modulations indicating that one could choose between these modulations depending on the application use case. Besides this, McNamara et al. [MBG09] conducted a similar second experiment using distractors in form of extra modulations which did not highlight a target. They placed these distractors in random distant spots in the image. Their results suggested a generally higher search performance correlation than the first study. They also compared the four modulation conditions with each other and solely found a performance difference between all three modulation conditions and no modulation at all. In addition, the distractors increased the performance compared to subtle and obvious modulations. According to McNamara et al. [MBG09], these distractors direct the gaze towards them resulting in a more complete gaze distribution over the image. In addition, the required refocused attention might avoid inattentional blindness. Hence, McNamara et al. [MBG09] conclude that the presence of distractors motivate the participants to consider the entire image. Although they considered 2D images instead of virtual environments, our approach using static distractors is quite similar as we also use similar distractors that look like our target.

Besides the presented work for static distractors, we did not find any further work studying the effect of dynamic or moving distractor objects on visual attention guidance techniques. Although Lange et al. [LSGB20] integrated some virtual wind within their environment, they did not measure or evaluate any potential implications regarding the wind. As previously mentioned in Section 2.3, Pinto et al. [POT08] investigated the influence of dynamic distractors on searching for certain targets among dynamic distractors in a 2D setup. For this, they used moving (i.e., bouncing from side to side) and blinking (i.e., switch a stimulus on and off) distractors. Several authors of visual attention techniques claim that motion is an effective visual feature to direct the attention to certain areas of interest [LSGB20; WH04; WLB+14]. However, it is possible that exactly this property may disturb or distract the effect of a technique, especially for techniques that exploit this feature [LSGB20; WLB+14]. Due to this, we investigate the effect of moving objects on visual attention techniques.

3 Implementation

This chapter provides an overview of our selected software components and the implementation of our study application. For the latter, we present our chosen environments and level of distractors which we simply refer to as distractorlevel for the remainder of this work. We also describe the selected techniques for our user study and how we implemented them. Afterwards, we also present the implementation of the general study application (e.g., the UI and measurements).

3.1 Selected Technologies

For the implementation of our virtual environment, we used Unity[1] and the Mixed Reality Toolkit[2] (MRTK) for immersive MR headsets. We shortly describe both components in the following.

Unity We built our entire study application with the game engine Unity. In general, Unity can be used to create 2D or 3D games but also allows to develop software for AR and VR [Teca]. For the implementation of immersive applications, it supports several additional platforms such as the Oculus, Microsoft Hololens or Windows Mixed Reality (e.g., by using the MRTK). We used the Unity Version 2021.2.8f1 as it was the latest version when we started our implementation. While Unity natively provides several basic geometric objects and features, it is also possible to add own components and behaviours using C# scripts.

Mixed Reality Toolkit for immersive MR Headsets We used the MRTK for setting up our VR-application and used the version 2.7.3.0 [Micb]. The MRTK is a project of Microsoft to provide several general components and features for the development of AR and VR applications in Unity. For this, it supports the Microsoft Hololens and Windows Mixed Reality Headsets (among them is our used HP Reverb G1). Besides the general setup for developing such applications, the MRTK also provides various basic UI and interaction elements and behaviour scripts that can be integrated into the own application.

[1]Unity Homepage: https://unity.com/

[2]MRTK Documentation: https://docs.microsoft.com/de-de/windows/mixed-reality/mrtk-unity/?view=mrtkunity-2021-05

3.2 Virtual Environments

Our study application consisted of two virtual environments, the so-called lobby environment and the task environment. In the following, we shortly describe both and also present their implementations. In general, we used multiple Unity scenes for the virtual environments. A Unity scene is an asset including certain parts or the entire implemented application [Teca]. For example, one can create a scene for each game level which has a particular surrounding environment, scene objects and behaviours. We decided for one scene for the lobby environment and three (similar looking) scenes of the task environment for each technique.

Lobby Environment The lobby environment was a regular office space within an office prefab of the 'Simple Office' [Dem] assets from the Unity Asset Store[3]. We decided for a certain room of the office prefab and also used its included interior. Additionally, we customized and re-arranged the interior to our requirements (see Figure 3.1). Since the participants did not see anything else besides the office room, we deleted several parts of the original prefab. However, we kept the hallway and the neighbouring office spaces that were visible from our office space to have a more natural and realistic appearance. We decided for a neutral starting environment due to a rather familiar appearance and to avoid potential learning effects or hints for the task. In general, the lobby environment was the starting point of the virtual study application. The participants conducted the training task and started the task environment for each technique within this environment. They also returned there after finishing all conditions (distractorlevel) for each technique.

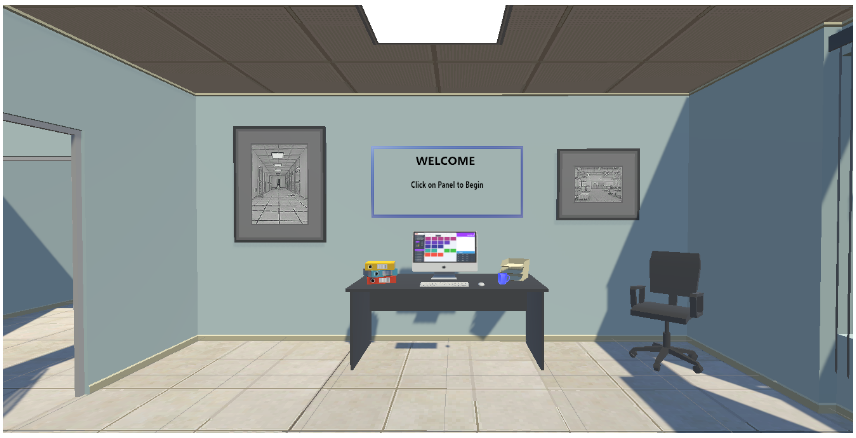


Figure 3.1: Lobby environment with welcome panel and training task.

Task Environment Since we wanted to extend the presented user study of Lange et al. [LSGB20] to some degree, we decided to use a similar surrounding and targets (see Figure 3.2). For this reason, our task environment looked like the forest environment of Lange et al. [LSGB20] and we also used the 'Winridge City' [Tecb] assets from the Unity Asset Store. It is important to note that these assets are currently outdated, but we purchased them earlier and hence, were still able

[3]Unity Asset Store: https://assetstore.unity.com/

to use them in our application. In general, we only integrated the terrain of their prefab scene for an immersive background. Besides the terrain, the task environment always consisted of a tree object for our target tree (using the assets ‘Realistic Tree 9’ [Gamb] of the Unity Asset Store). We also added several red apples as potential target objects which had a unique name for selection and evaluation processes within the task (see Section 3.5). Additionally, to make them clickable, we integrated an ‘interactable’-component provided by the MRTK. Depending on the distractorlevel (see Section 3.4), there were several other distractor-based elements within the scene. After starting each technique in the lobby environment, the participants stayed within our task environment scenes until they finished all distractorlevel.

Figure 3.2: Task environment with first panel and no distractors.

3.3 Selected Visual Attention Techniques

Since we aim to compare visual attention techniques, we selected the following three. We summarise these techniques and their previous research findings. Furthermore, we provide an overview on how we implemented them in our study application.

3.3.1 Deadeye

Our first technique is Deadeye by Krekhov and Krüger [KK19] which renders a target object to only one eye while the other eye does not perceive the target at all. According to Krekhov and Krüger [KK19], Deadeye is a beneficial technique because it exploits the binocular rivalry (see Section 2.4) and does not alter any visual features (e.g., by recolouring or using motion) of a target object.

Within their work, they first evaluated whether Deadeye is perceived preattentive and can be perceived with other visual features (i.e., conjunction search properties) using a 3D TV and shutter glasses. For this, they conducted a user study by showing participants several images with a varying amount of distractors. Based on the general study setup for preattentive features (see Section 2.2, [HE12]), they decided for measuring the performance based on accuracy. Hence, after a training task, their participants had to report the presence or absence of a target (i.e., a circle highlighted by Deadeye) among several distractors (i.e., several circles without highlighting). According to their results, Krekhov and Krüger [KK19] concluded that Deadeye can be perceived preattentive as their accuracy did not significantly differ compared to common visual features and regarding a varying count of distractors. Their results also suggested that a target is rather missed than falsely perceived to be present and that objects in greater distance are harder to identify. Within their study, they also varied the rendering of the target between each eye. According to Krekhov and Krüger [KK19], their results showed that there is no difference between showing the target to the dominant or non-dominant eye. Furthermore, Deadeye is also perceivable when presenting it with other visual features such as colour. Based on their subjective feedback, they concluded that Deadeye did also not lead to any strains (e.g., headache).

In another work by Krekhov et al. [KCWK19], they also investigated whether Deadeye can still be perceived in a virtual environment using a HMD. For this, they used 3D cubes instead of 2D circles. While they only used homogeneous objects for their target and distractors for their first study, they extended their experiment by also investigating Deadeye using varying types of distractors (i.e., by colour, shape, 3D depth). Therefore, they conducted two experiments, the first being similar to their initial experiment with a virtual environment, 3D objects and only homogeneous distractors and a second experiment with heterogeneous distractors. The results of both experiments showed that Deadeye can still be perceived preattentive and that a virtual environment did not influence the performance. In addition, the distractor type did not have an effect on the performance of the participants. Lastly, Krekhov et al. [KCWK19] also investigated their proposed technique by conducting an exploratory study for applying it to volume rendering. According to Krekhov et al. [KCWK19] Deadeye can be applied for such visualisations even though they indicate that their participants still prefer colour scales.

To render a target for one eye but not for the other, we require a setup that allows a separate rendering and visualisation for each eye. In Unity, a camera usually presents the environment to the user and hence, there is only one single camera within the scene [Teca]. This means for a VR application that the camera represents the user in the virtual environment and hence, their field of view. For our setup, we subordinated two cameras instead of one camera to the parent camera object in the scene. As Unity requires exactly one main camera, we tagged the camera responsible for the left eye to be our main camera. Within the camera properties of Unity, one can adjust the target eye to which the camera renders the scene. Usually this is set to both eyes, but in our setup we respectively set it to either the left or the right eye. In addition, we added to both cameras a tracked pose driver to be able to view the scene using a HMD. The camera settings also include a culling mask which is responsible to render specific layers (certain parts of scene objects) selectively. By default, a camera renders all layers and every scene object is assigned to the default layer [Teca]. However, it is possible to exclude (or include) layers from the culling mask and assign scene objects to predefined or custom layers.

We created an own layer, called ‘left eye layer’ which was only assigned to the current target object. While the camera for the left eye rendered all layers, we excluded the ‘left eye layer’ from the culling mask of the camera for the right eye. Thus, the target apple object is only visible for the left eye but not for the right eye (see Figure 3.3). We did not switch this representation randomly or based on the participant’s respective eye dominance (e.g., show it only to the dominant or non-dominant eye) due to the following reasons. Firstly, as mentioned before, Krekhov and Krüger [KK19] did not find any significant differences between rendering the target to the participants dominant or non-dominant eye, indicating that it does not have an effect which eye perceives the target object. In addition, as we wanted to extend the study design of Lange et al. [LSGB20], we tried to keep their settings. Lange et al. [LSGB20] did also not change the rendering setting based on the eye dominance.

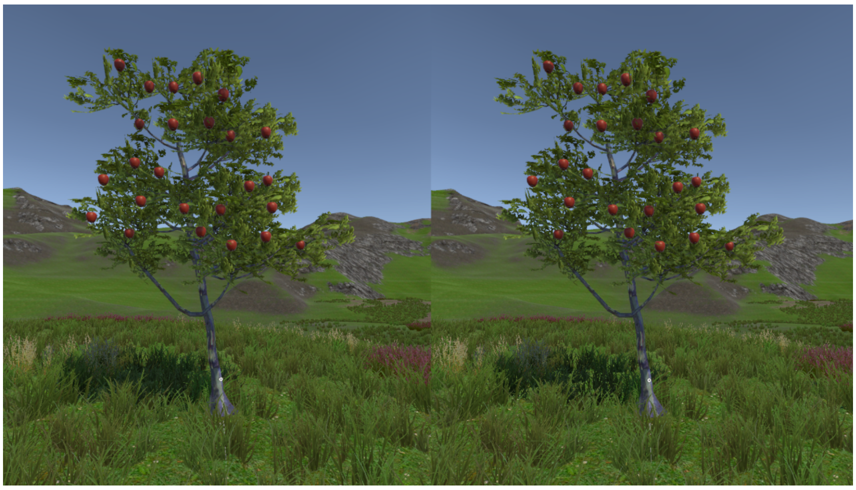


Figure 3.3: Deadeye with no distractors.

3.3.2 HiveFive

Lange et al. [LSGB20] presented and evaluated a technique called HiveFive. It is inspired by bees which mostly occur in swarms and can be integrated in several environments. For their study but also for our work, they used several yellow, basic spheres that move in swarm-based behaviour. They decided for the colour yellow as according to them, it can be recognized well in the periphery. Lange et al. [LSGB20] conducted two studies to evaluate their technique. Within the first study, they compared the visual perception of a non-biological swarm motion (using the flocking algorithm by Reynolds [Rey87]) and a biological swarm motion (using a dataset with recordings of swarm trajectories). Their results showed that the non-biological swarm motion performed better as it was perceived faster. Hence, they applied the non-biological swarm for their second user study.

In the second user study, Lange et al. [LSGB20] studied the performance of their technique with four other visual attention guidance techniques using a within-subject approach. They selected the Subtle Gaze Direction (see Section 2.4), Blurring (renders only target location sharp) and Arrow (a yellow arrow pointing towards the target) and similar to our work, Deadeye. They used a forest environment and an apple tree with several apples as target objects. After a training task, participants had to follow a sphere until a technique highlighted a certain apple. As soon as the participants perceived the technique, they had to press a button. By pressing the button, the technique disappeared and the participants had to select the previously highlighted target within five seconds. A trial was stopped if participants did not select an apple within this interval or did not perceive the technique so that participants did not select random apples. According to Lange et al. [LSGB20], HiveFive is among the best techniques in terms of response time and accuracy. In addition, they highlight that HiveFive seems to negatively affect the perceived sense of presence the least and preserves immersion.

Given that the swarm of Lange et al. [LSGB20] consisted of several small, yellow basic Unity spheres, we created a similar scene object fitting these requirements. To make our study as comparable as possible, we contacted the authors for further insights in addition to consulting the paper and the supplemental material. We adjusted the swarm parameters of the colour, the shader and the scale of our spheres accordingly after the pilot study.

Lange et al. [LSGB20] based their swarm behaviour on the flocking algorithm by Reynolds [Rey87]. Before receiving detailed information from the authors which was missing from the supplemental material, we needed a custom solution for the parameters. We reimplemented a free online tutorial by GameDevChef [Gama]. Their implementation includes several parameters for cohesion, avoidance and alignment as well as the flocksize, bound and obstacle distance which can be customized based on the requirements. Based on the described parameters by Lange et al. [LSGB20] and self-empirical testing, we adjusted the swarm behaviour to a reasonable appearance surrounding the target apple. Given that our custom solution and the attached solution by the authors substantially differed but provided comparable results, we stayed with our solution. The final representation of our swarm can be seen in Figure 3.4.

Figure 3.4: HiveFive with no distractors.

3.3.3 Circle

Besides the rather subtle binocular disparity-based technique Deadeye and the fast motion-based technique HiveFive, we intended to additionally use a static and fairly common approach for our study. As previously mentioned (see Section 2.3), a common approach is applying preattentive features for attention guidance and also the combination of more than one feature can efficiently direct the attention to a certain target [WH04; WU18]. As colour and shape are rather common preattentive features, we decided to use a combination of both as a yellow circle surrounding the target object as an additional technique which can be seen in Figure 3.5. We decided for a yellow colour since Lange et al. [LSGB20] claimed that it is well perceived in peripheral regions and hence, applied it for the Arrow technique in their user study. Based on the provided supplemental material of Lange et al. [LSGB20], we modified the material properties (e.g., colour and shader) of the Circle to be similar to the Arrow.

For the circular shape, we used a free ring object of the 'Simple Collectibles Pack' [Wal] of the Unity Asset Store. We decided for a 3D object (ring) instead of a 2D object (circle) because all other scene objects are three-dimensional and it is more natural appealing. In order to highlight a target object, the Circle had to surround the target object. Therefore, we increased the radius of the ring object such that it almost perfectly enclosed the target object but was still fully visible. When selecting a target object, the Circle centre shifted to the next target object position.

Figure 3.5: Circle with no distractors.

3.4 Distractors and Selected Levels

Based on our aim to investigate the influence of distractors on visual attention techniques, we decided for three different, increasing distractorlevel. This section provides an overview of their general structure and appearance. In addition, we give some further details about their implementation.

In general, we subordinated all distractor objects for a respective distractorlevel to a generic parent object (e.g., a static distractorlevel parent object with its child static distractor objects) to facilitate the access within the application. Therefore, it is possible to set the parent object active to make all subordinated distractor objects and their behaviours visible instead of doing it one by one. The only exception was the target tree as it stays visible in all distractorlevels. Due to consistency reasons, we also integrated a parent object for no distractors even though it did not have any subordinated objects. We used a dictionary (C# equivalent of a map [Mica]) to store for each technique the respective distractorlevel order so that it could be accessed and set the right distractor objects active depending on the current distractorlevel.

3.4.1 No Distractorlevel

The so-called no distractorlevel did not have any distractor objects at all (see Figure 3.6). Hence, it looked like the previously explained task environment (see Section 3.2) without any additional objects.

Figure 3.6: No distractorlevel.

3.4.2 Static Distractorlevel

The static distractorlevel was the second level and consisted (besides the target object) of additional static distracting objects (see Figure 3.7). This means that these objects did not move and we did not integrate any additional motion. Like McNamara et al. [MBG09] and Olk et al. [ODZK18], some of our static distractors were similar to the target object. Hence, we integrated several trees with apples and placed them around the target tree. For the tree and apples, we used the same assets as for the target tree and apples (see Section 3.2, [Gamb]). However, we also added some flower beds next to the target tree using the 'Ornamental Flower Set' [Stu] of the Unity Asset Store.

Figure 3.7: Static distractorlevel.

3.4.3 Dynamic Distractorlevel

Besides the static distracting objects, we also included some additional moving objects and effects for the dynamic distractorlevel (see Figure 3.8). Firstly, we used a bird swarm (using 'Living Birds' [din]) and a bee swarm (using 'FantasyBee' [amu]) that behaved similar to the HiveFive swarm using our implemented solution of the flocking algorithm by Reynolds [Rey87]. For both swarms we respectively adjusted the amount of objects and the parameters of the algorithm so that they were always in the nearest field of the participants but also acted rather more naturally by flying greater distances. Since Olk et al. [ODZK18] claim that similar features of static targets and distractors might influence their perception, we assumed a similar effect for moving targets and distractors.

Due to the variety of materials for different bird objects of our assets, we decided to add another bird swarm. For these birds, we created a script that spawned a random number of three to five birds that flew randomly one of three predefined paths. To create these paths, we used a Bezier Path Creator [Lag], provided in the Unity Asset Store, to draw them into the environment. Since birds are one of the most common moving objects in nature, we also integrated smaller and distant birds that flew a bit higher than the rest of the birds. For this, we placed an integrated particle system of Unity in small distance to the target tree and changed the particle appearance to a bird prefab using the same bird assets.

As we had three different flower beds, we decided to add some motion for them as well. Hence, we added for two flower beds two distinct swarms of butterflies using the ‘Butterfly (Animated)’ [Gmb] from the Unity Asset Store and let them move by using their integrated animation. For the third flower bed, we used another particle system to present them as pollen by changing their appearance to be white and slightly transparent. Similar to Lange et al. [LSGB20], we also added some wind to move the limbs and leaves of the tree. Lastly, we added some physical behaviour (i.e., gravity) to a random apple of the right tree next to the target tree. Due to this physical component the apple fell down on the grass. Due to our study setup, we delayed the process instead of directly executing it when the dynamic distractorlevel was present.

Figure 3.8: Dynamic distractorlevel.

3.5 Virtual Study Setup

This section provides a short overview of our implementation of the components for the virtual setup as well as the study procedure within the virtual environment. As mentioned in Section 3.2, we used multiple scenes that we connected by scripts to realise behaviour and associated interactions of the participants. Besides the respective technique, each task environment scene integrated the same scene objects, behaviours and scripts. The scripts of the lobby environment scene controlled the whole study procedure and started all necessary processes (virtual experience in general and for all techniques). Our main components consisted of scene objects and attached scripts. While the scene objects structured the appearance of the scene content, the scripts handled the behaviour and processes. We tried to handle the majority of our study procedure by scripts so that we only had to increment the participant ID during the study conduction.

3.5.1 Technical Details

It is important to note that Unity destroys all scene objects and references of a scene when loading or starting another scene. In our case, the lobby environment scene is responsible for the whole study execution meaning that as soon as a participant transfers to the task environment scene for a technique, by default Unity deletes all participant information of the current status (e.g., distractorlevel order for the respective technique or the completion of the training task). There would be also no access any more to custom adjustments such as the user position and rotation when the participant re-enters the lobby environment scene after finishing all conditions (distractorlevel) for a certain technique. Since all this information is necessary for a successful conduction of the study, we wanted to preserve this information beyond scenes. For this, Unity provides two solutions. The first solution is keeping an entire scene awake while being within another scene. However, this caused in our case several problems regarding the camera count within the scene. We decided for the second method to keep only certain aspects (e.g., scene objects) awake. With that we kept the scene object with the attached study procedure behaviour script awake and when re-entering the lobby environment scene it handled the necessary aspects (e.g., no displayed training task elements or displaying the right panel). Hence, it did not interrupt any processes within the task environment scenes. Additionally, we added a script which is always executed when re-entering the lobby environment scene to update information, previous positions and rotations of the participant.

3.5.2 User Interface

We also added a user interface consisting of several panels of the MRTK to inform about the next steps. Our panels were quadratic, transparent signs surrounded by a coloured border that performed some interactions by clicking on them. Since they had to instruct the participants in some way, we added text depending on its purpose. The clickable behaviour is realised by the 'interactable'-component of the MRTK attached to the panel. Due to this component, the panel notices a click within its borders and triggers the customized, assigned function.

For the lobby environment scene, we had the following panels that triggered the subsequent behaviour. The first panel within the scene welcomed the participants within the virtual environment and can be seen in Figure 3.1. By clicking on it, it disappeared and the training task panel replaced it. This panel did not trigger any specific behaviour itself but instead explained the training task (i.e., nothing happened by clicking on it). After the successful completion, it automatically changed to the task panel. The task panel was the last panel within the lobby environment scene before transferring to the task environment scenes depending on the respective technique. Hence, by clicking on this panel, the task environment appeared with the first distractorlevel present. In general, it loaded the scene depending on an index that looked up the next technique in the technique order list. As soon as the participants finished all distractorlevel, they returned to the lobby environment scene, seeing the same task panel again to transfer again to the task scene environment for the next technique. All in all, the participants had to click on this task panel three times for three techniques. However, when the participants returned to the lobby environment scene after completing the tasks for all three techniques, the end panel occurred. Similar to the welcome panel, it noted that the participant had finished all techniques or tasks and thanked for the participation. For that, we used a different material to indicate a different behaviour as it stated the end of the virtual experience. We decided that the participant had to return to the lobby environment scene to have a more consistent ending. Additionally, the successful return to the lobby environment scene marked a successful completion of our study setup (i.e., a reassurance that it saved the participant's data).

Besides the lobby environment scene, the task environment scenes owned the following two panels. The first panel occurred when being the first time in the task environment scene for a certain technique and in general before highlighting the first target object for each distractorlevel. While the panel was active, the respective distractorlevel could be considered without any task or time pressure. The panel also suggested the participant to take a moment within the environment before starting the task by clicking on the panel (see Figure 3.2). Once clicked, it disappeared and the task started meaning that the first target object was highlighted and the respective measures were measured. We decided for this panel to give the participants time to accustom within the new environment and to ask potential open questions (especially for the first technique). In addition, integrating such a panel can avoid potential measurement biases due to longer loading times caused by the hardware. Thus, we assumed that as soon as the participant clicks on the panel, they are confident enough to do the task. The second panel appeared after finishing the task for one distractorlevel. It suggested to put the HMD off and to fill out the questionnaire before starting the next distractorlevel. While being active, only our plain task environment was visible without any distractors so that participants could start the next distractorlevel when feeling ready. By clicking it, the next distractorlevel was visible and again the first panel was shown to start the next task. However, the second panel was not active if the participant finished the last distractorlevel but instead, they directly returned to the lobby environment scene to fill out the questionnaire and potentially start the next task scene environment for the next technique.

3.5.3 Training Task

We decided for a mandatory training task to ensure that all participants understood how to click on a virtual object to avoid potential hardware struggles within the study task. Especially due to our measured completion time, we wanted to exclude such biases that potentially influence our results. Since we did not want to give any hints about the highlighting techniques, we used a general clicking task without any highlighting and also did not measure anything. The participants performed the training task within the lobby environment scene to also not give any hints about the task environment. The training task fitted thematically into the surrounding to preserve immersion. It consisted of several common office objects (e.g., folders, monitor or a cup) lying on a desk (see Figure 3.1) using prefabs provided by the 'Simple Office' [Dem] assets of the Unity Asset Store and the MRTK. They differed in size and were clickable using again the 'interactable'-component by the MRTK. If a participant successfully clicked on an object, it disappeared (i.e., set the scene object inactive).

3.5.4 Study Task

As mentioned before, we extended to some degree the work of Lange et al. [LSGB20] to evaluate the influence of additional distractors. Hence, we decided to have similar a structure of highlighting target objects with our selected techniques by also using a tree in front of the user with various red apples as potential target objects. However, we decided to use a different task procedure for comparing the performance of the attention guidance techniques. In our task, the participants only had to find the highlighted apple in front of them while being confronted with different distractors depending on the current level. As soon as they found it, they had to click on the specific apple object. After successfully clicking on the right apple, either another apple was highlighted or in case the predefined amount of apples were found per distractorlevel, the respective UI panel informed about the following steps.

For selecting a certain apple object of the target tree as a target, we subordinated all potential target apples to a parent object which knew the amount of its child objects. In addition, we implemented a function which chose a random number in between one and the total amount of target apples. The apple with the corresponding number was the target, since the name of each target apple consisted of the string ‘targetapple’ and a unique number. We also saved all names of the chosen target apples for the evaluation. After selecting a target, it had to be highlighted. According to the current scene name (given that it represented the current technique), the program executed the respective function that highlighted the selected target respectively. The explanation on how we implemented the technique can be found in Section 3.3. As mentioned before (see Section 3.2) all target apples were clickable. By clicking on an apple, its appended script returned the unique index of the apple. Hence, it was possible to verify if the currently clicked apple was the target or not by comparing the current name and the target name. Depending on the result, either it returned a match or a mismatch. For a mismatch, it increased the wrong clicked object index which we measured besides the completion time (explained in more detail in the next subsection). If it was the right apple objects, it stopped measuring the completion time and either selected, highlighted another target, ended the distractorlevel (i.e., found all targets for one distractorlevel, presenting the second task environment panel) or ended the technique (i.e., found all targets of all distractorlevels for this technique, returning to the lobby environment scene).

3.5.5 Measurements

We counterbalanced the order of the techniques using a Latin square design and randomised the order of the distractorlevel. For the former we developed an automatic process by creating in advance a CSV file with 30 rows including the participant ID and the counterbalanced order of the three techniques. By starting our study application, a script read the row based on the current participant ID and stored the respective order in a predefined array of size three. This array is accessible whenever the technique order is required. Moreover, a script also pre-calculated the random order of the distractorlevel after storing the technique order. Therefore, it was only necessary to access the associated list (depending on the technique) instead of calculating it when starting the task environment scenes. It stored the randomised list of the distractorlevels in a dictionary with the technique as a key.

We measured two variables within the virtual environment, the completion time and the wrong clicked objects per target object. For the completion time, we saved the time difference between the start of highlighting the target object and a successful click on the right target object. Due to the unique naming of the target objects we were able to compare the name of the current clicked object and the current target object. When finding a mismatch, we increased the index of wrong clicked objects by one. For the wrong clicked objects, we only took other apple objects of our target tree into account and not other visible objects in the virtual environment (e.g., apples of other trees, flower beds) because it would bias the data values as such objects were not visible for the no distractorlevel. We stored both measures for each target in two distinct lists. Similar to the randomised distractorlevels, we created a dictionary for the completion time and the wrong clicked objects that included all values of all distractorlevels for all techniques again using the respective technique as a key. Furthermore, we stored a list and respectively, a dictionary, for all random chosen target object names.

To save all the above described data for the evaluation, we created a CSV file by a script. After the participants finished all techniques and returned to the lobby environment scene, the program automatically appended the data of the current participant in the file. It saved the respective participant ID, the order of the techniques and distractorlevels as well as the target apple name and its corresponding completion time and wrong clicked objects.

4 User Study

This chapter provides an overview of our user study design. We first present our study goal as well as our hypotheses and measured variables. In addition, we summarise our study procedure as well as the conducted pilot study and its resulting implications, modifications for our study design. In the end we present the results of our demographic questionnaire, meaning the demographic information of our participants of the user study.

4.1 Study Goal

As previously discussed in Chapter 2, the majority of research in attention guidance only evaluated attention techniques in VR using realistic environments but did not include any other, potentially disturbing, objects. In addition, most researchers try to avoid common 2D techniques such as preattentive features to preserve the immersion and to enhance the sense of presence within VR. However, for an immersive environment, it is necessary to evaluate visual attention techniques in more realistic and crowded environments since it is rather uncommon to have a calm surrounding without any interaction or disturbance in real life. Since CVR is also a huge application area of VR and attention guidance, it is even more important to have a strong and efficient attention guidance as there are additional entities present [NMH+16; RBH19].

In this work, our aim is to evaluate the influence of additional entities (e.g., similar looking non-moving objects and moving objects) as distractors on attention guidance techniques. For this, we compare our selected techniques (see Section 3.3) Deadeye, HiveFive and Circle. Therefore, we extend the work of Lange et al. [LSGB20] by using a similar setup as well as the surrounding environment, but include an adjusted task and further conditions using static and dynamic distractors.

4.2 Research Questions and Hypotheses

As this work aims to provide further insights for applying visual attention techniques in realistic contexts without any physical or psychological limitations, we defined the following research questions:

- Are participants able to identify target objects that are highlighted by the selected techniques?
- Do static distractors influence the accuracy of identification for identifying the target?
- Do dynamic distractors influence the accuracy of identification for identifying the target?
- Are there any changes in mental or physical load according to techniques or distractor differences?

According to these research questions, we decided for the following hypotheses:

- H1: The completion time increases for all techniques with increasing distractorlevel.
- H2: The completion time increases more for HiveFive with dynamic distractors than with other techniques.
- H3: Deadeye has higher mental demand and effort compared to other techniques with increasing distractorlevel.

4.3 Apparatus

The study was conducted within the VR Laboratory of the Institute of Visualisation and Interactive Systems at the University of Stuttgart. We followed the hygiene measures established at that time due to the Covid-19 pandemic. The main part of our user study was within VR. For this, we used the implemented virtual environments which we described in Section 3.2. We used the Windows Mixed Reality Headset HP Reverb G1 and their associated motion controllers. In addition, we used the online survey tool Limesurvey[1] to create and conduct our questionnaires. We presented them on a PC and a standard 2D display using Windows 10.

4.4 Methods and Procedure

We decided for a within-subject approach such that each participant had to perform all conditions (3 techniques x 3 distractorlevel = 9 conditions). In the following, we describe our independent and dependent variables. In addition, we present the conducted procedure of our user study (and pilot study).

4.4.1 Independent Variables

We had the following two independent variables:

- Technique (Deadeye vs. HiveFive vs. Circle)
- Distractorlevel (No vs. Static vs. Dynamic)

We described both variables in Section 3.3 (technique) and Section 3.4 (distractorlevel) in further detail. As previously mentioned, we counterbalanced the techniques using a Latin square design. We had for each technique three different distractorlevel, resulting in 9 conditions per technique and distractorlevel. The order of the presented distractorlevel of each technique was randomised for every participant. Our implementation (see Section 3.5) randomly selected 10 targets for each distractorlevel, resulting in 30 targets for each technique.

[1]Limesurvey Homepage: https://www.limesurvey.org/de/

4.4.2 Dependent Variables

We measured the following **two quantitative measures** for each target within the virtual environment as dependent variables. Further details on the implementation are described in Section 3.5. In general, as we measured both values for each target, it resulted in 90 values (10 per distractorlevel, 30 per technique) per participant.

- Completion Time (difference between highlighting and a successful click on the right target)
- Wrong Clicked Objects (count of incorrectly clicked apples per target)

In addition, participants had to fill out a questionnaire after each distractorlevel. The questionnaire consisted of a standard **NASA TLX**[2], some own questions and a further text field for further comments. The NASA TLX consisted of the following subscales to measure the subjective perceived workload for each distractorlevel of every technique. We used a scale ranging from 0 (indicating very low, for performance failure) to 20 (indicating very high, for performance perfect).

- Mental Demand
- Physical Demand
- Temporal Demand
- Performance
- Effort
- Frustration

For the following **own questions**, we oriented our formulation based on the own questions of Lange et al. [LSGB20] and Krekhov and Krüger [KK19]. We used 5-Likert Scale (strongly disagree, disagree, neutral, agree, strongly agree).

- I could perceive highlighted object well.
- The technique fits well into the overall scene setting.
- The current visual scene elements distracted me from the target search.
- The current visual scene elements made the task more difficult.

The first one should indicate a rather general impression whether participants could perceive a target when being highlighted by a certain technique. For the second, we expected that certain techniques would fit more in an environment due to their immersion preserving properties. The last two statements dealt more with the impact of the distractors and their potential of disturbing the participant. We assumed that a higher level of distractors would disturb and distract more.

[2]NASA TLX Homepage: https://humansystems.arc.nasa.gov/groups/tlx/

4.4.3 Procedure

Before the participant arrived, we sanitised the HMD as well as all necessary entities of our study (e.g., keyboard, table). After informing the participants about the current hygiene measures, the participant could sign the consent form and the privacy form. Then, they had to fill out the demographic and a visual disability questionnaire to test for a potential red-green weakness. The demographic questionnaire consisted of general questions about the participants (age, gender), their general visual abilities (short-, long-sighted as well as other known visual disabilities) as well as their dominant eye. In addition, we added a 5-Likert Scale to rate their VR experience and their (PC-) gaming usage. In case they did not know their dominant eye, we informed them about the standard test for eye dominance. The visual disability questionnaire added five Ishihara tables which consisted of four tables having a number and one did not have any (for people with normal vision). They had to state the perceived number (or ‘no’ in case if they did not perceive any number) in a free text field. Subsequently, they got some general instruction about the HMD and its motion controller as well as an introduction of the further procedure and their task (see Section 3.5). We held the explanation of the study rather general by only pointing out that an object would be highlighted or popping out instead of explaining how and if they would identify an highlighting object, they should click on it. In general, we also mentioned that everything would be within the bigger field of view, meaning that they would not have to move or look behind. However, it could be outside in the periphery of the field of view. We also noted that they should click very cautious because of the motion controller and to stretch or lift the arm a bit. We also showed them the lobby environment in advance so that they were familiar what they would perceive when putting on the headset.

After answering potential questions, participant were able to put on the HMD. Before starting our Unity application, we gave them a few moments within the default scene of the Windows Mixed Reality Headsets to accustom themselves within a VR environment using motion controllers. Within the lobby environment, we adjusted their position so that participants had comparable distances to the objects and they could start the training task (see Section 3.5). After completing this, they could start the task environment for the first technique and look at the first distractorlevel. When there were no open questions, they could start the task and try to find the first 10 targets. While the participants were completing the task, we noted within the questionnaire which technique and distractorlevel the participants had to do. After successfully finishing a distractorlevel, the participants took off the HMD and filled out the questionnaire using the PC and keyboard considering the NASA TLX, own questions and the further comment text field. After they filled out the questionnaire, they could continue with the next two distractorlevels. After finishing all three distractorlevels, they returned to the lobby environment and also took off the HMD to fill out the last questionnaire for this technique. The remaining two techniques had the same procedure to the start task environment by clicking on a panel, passing through all three distractorlevels and questionnaires and returning back to the lobby environment. After completing all three techniques with all including questionnaires, they had to fill out the final questionnaire. The final questionnaire consisted of a 5-Likert Scale on physical and mental strains, a ranking of the techniques including a text field to give reasons for the chosen ranking for every technique. As for the study questionnaire, we also included a further comments text field for final comments or feedback. During this process, the program automatically saved the measured data for completion time and wrong clicked objects. In the end the participants got a compensation of 10€ and had to sign a form that they received the reward. All in all, the whole study procedure took approximately 60 minutes per participant.

4.5 Pilot Study

Before conducting our user study, we wanted to test our application and study procedure to find potential problems and ambiguities and to estimate the entire study procedure. Hence, we conducted a pilot study after the finalisation of our implementation and could apply the previously explained procedure using also a first draft of our questionnaires. However, participants were not compensated at the end of the study. We had 5 participants (2 females and 3 males) for our pilot study who were all actively recruited at the Institute of Visualisation and Interactive Systems at the University of Stuttgart. Overall, it took 60 to 90 minutes to complete the study, depending on the level of detail of the given feedback. In the following, we first present a broad review of the results. Subsequently, we explain our modifications based on the feedback regarding the implementation and the questionnaire.

Since we only had five participants, we did not evaluate the measures using statistical tests. However, we considered the mean completion time for each distractorlevel of every technique. We saw a slight trend for the completion time regarding Deadeye having the highest mean and Circle having the smallest mean but this trend was very small. Considering the wrong clicked objects, we only had values for Deadeye. For the questionnaire, we only looked at the data considering the further comments text field. For Deadeye, the participants described it as unnatural and it seemed more like a vision or hardware mistake. In addition, it was difficult to perceive in peripheral regions. In general, most of the participants mentioned that the first (few) targets were really complicated but after knowing the effect they felt more confident and found the targets faster. One participant mentioned that they had to browse through all apples to find the one appearing differently. The Circle appeared for some of the participants more foreign in the scene and was in general easy to distinguish in the dynamic scene due to its colour and static appearance. HiveFive appeared nice like bees, but with real bees more difficult to get. A participant mentioned that the scene included partly too much movement.

We did not find any major bugs as our study conduction did not fail at any point. There were only some comments for the panels which we later adapted (e.g., unambiguous formulations). As mentioned before (see Section 3.3) we adjusted some parameters of our swarm and flock algorithm after the pilot study. We also changed the structure of our CSV file (for the completion time and wrong clicked objects) for an easier evaluation after the user study and added a column for the chosen target apple names.

Within our first draft of the demographic questionnaire we did not ask about the (PC-) gaming usage. As one of our participants stated that they could mask out the distractors due to their regular consume of video games, we decided to add this to the 5-Likert Scale besides the VR experience. Our first formulation of the second own question (‘was well embedded in environment’) resulted in several questions about the meaning of ‘embedded’. For this reason, we re-formulated it to its final version (‘fits well into overall scene setting’). Since several participants commented their most and least favourite technique, we decided to add a ranking of the techniques (and a text field for its reasons).

4.6 Participants

We had 30 participants (9 female and 21 male) in our user study. The age of the participants varied between 18 and 55 years. The specific distribution can be seen in Figure 4.1 and shows that most participants were between 22 and 30 years old. Furthermore, 10 participants were short-sighted and 2 participants were long-sighted. Only one participant mentioned that they had a red-green weakness. For two out of the five Ishihara tables, all participants perceived the correct number. For another table, only one participant did not perceive the respective number correctly. Surprisingly, the remaining two had two instead of one wrong answers. However, one out of these two was the plate without any number. It could be possible that the participant was confused and thus, guessed a possible number. In general, we asked the participants whether they could distinguish between the scene elements and their background when being within the virtual task environment scenes for the first time before starting the highlighting task. Since all participants were able to, they all could perform the remaining study procedure. We had 10 participants with a left dominant eye and 20 participants with a right dominant eye. The results of the 5-Likert Scale regarding the VR experience and the (PC-) gaming usage are shown in Figure 4.2. While they had mostly less VR experience, most of the participants had a rather high (PC-) gaming usage.

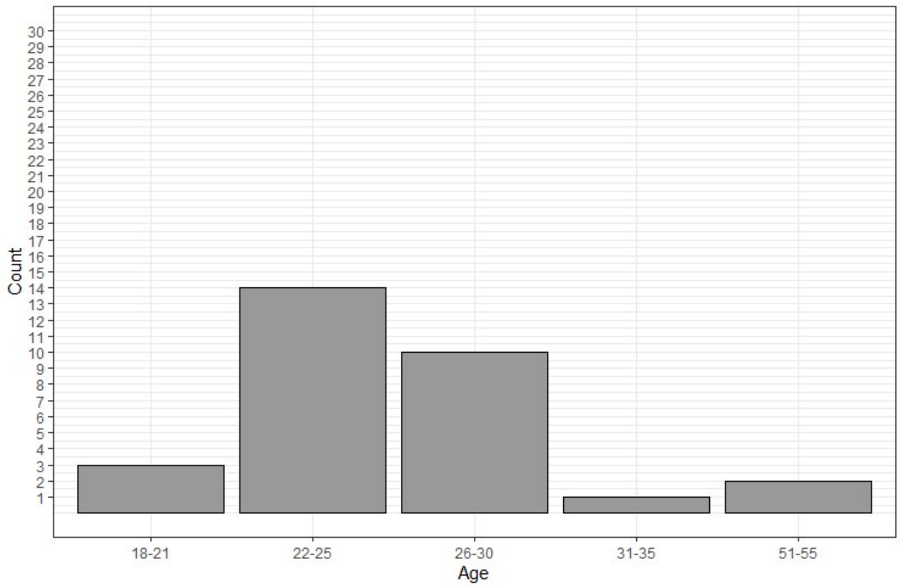


Figure 4.1: Age distribution of the study participants.

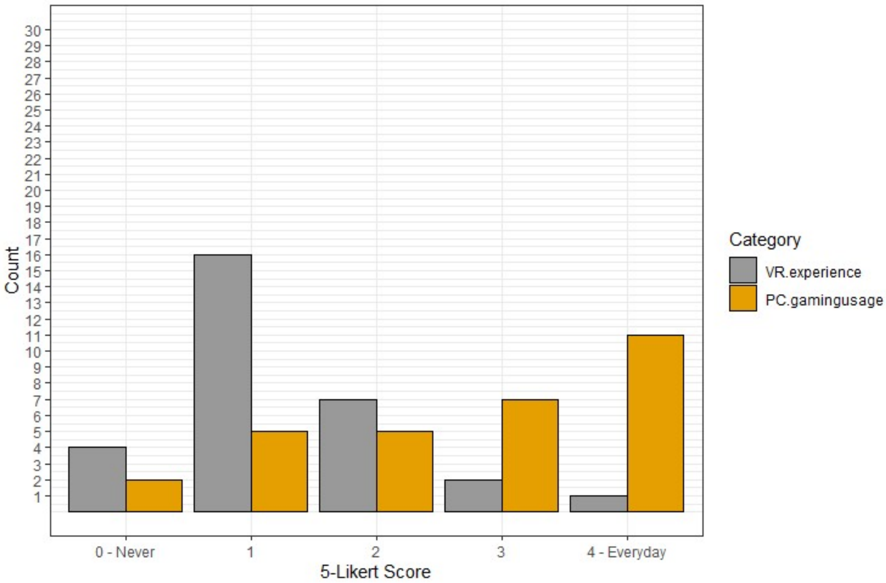


Figure 4.2: VR experience and PC-Gaming usage of the study participants.

5 Results and Discussion

This chapter includes the evaluation of the user study data. Using our collected data, we present and discuss the study hypotheses alongside other salient findings. In the first two sections we address findings of the completion time and wrong clicked objects which were measured during the task. Within the completion time, we also consider the results of H1 and H2. Afterwards, we look at the data of the NASA TLX and our own questions which the participants completed after each condition (3 techniques x 3 distractorlevel). As we also gathered qualitative data from the participants, we discuss this in the last section. In this case, we consider the further remarks of the participants that could be given after each condition as well as the perceived mental, physical strains and the technique ranking of the post study questionnaire. For our evaluation, we used RStudio[1], a free and open source IDE for R. Within this IDE it is possible to directly execute the code and it provides several tools for plotting [RSt]. R is a free and open source programming language for statistical computing and includes many tools for statistical methods [Fou].

5.1 Completion Time

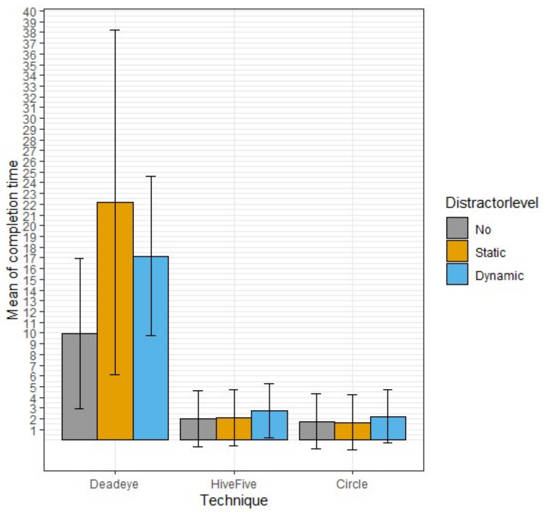
First, we report the results of the completion time of the targets which was measured while the participant performed the VR-application. We measured per target the time difference between highlighting and successfully clicking on it. Since we had 10 targets per distractorlevel, we had 30 values per technique and thus, 90 values per participant. To aggregate the completion time per participant, we calculated the mean of the 10 targets for one distractorlevel resulting in 9 values per participant. We again calculated the mean of the completion time for each distractorlevel of every technique leading to nine different values. These means (M) and standard deviations (SD) can be seen in Table 5.1. In addition, we plotted the mean and the confidence interval in Figure 5.1.

Distractorlevel	Deadeye	HiveFive	Circle
No	M= 9.94 (SD= 18.77)	M= 2.01 (SD= 6.95)	M= 1.77 (SD= 6.90)
Static	M= 22.15 (SD= 43.04)	M= 2.09 (SD= 6.99)	M= 1.67 (SD= 6.86)
Dynamic	M= 17.17 (SD= 19.92)	M= 2.77 (SD= 6.69)	M= 2.22 (SD= 6.64)

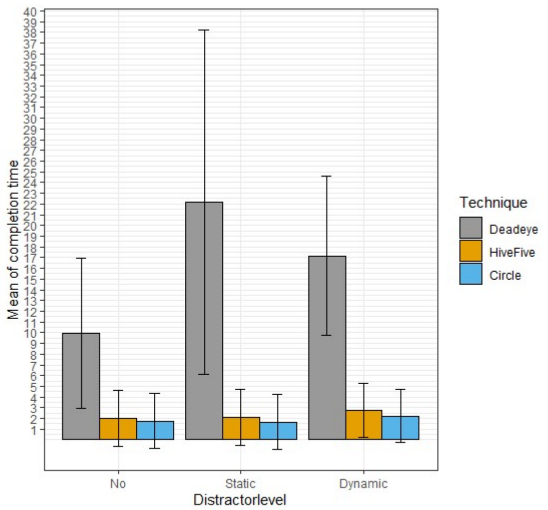
Table 5.1: Mean and SD of the completion time by technique and distractorlevel.

In both representations it can be seen that Deadeye has the highest completion time for all distractorlevel, while there is only a small difference between HiveFive and the Circle. In general, for static and dynamic distractorlevels the completion time is a bit smaller for the Circle than for HiveFive. While the completion time only increases for HiveFive the way we would expect it,

[1]RStudio Homepage: https://www.rstudio.com/



(a) Ordered by technique.



(b) Ordered by distractorlevel.

Figure 5.1: Error bars showing the 95% confidence interval for the completion time.

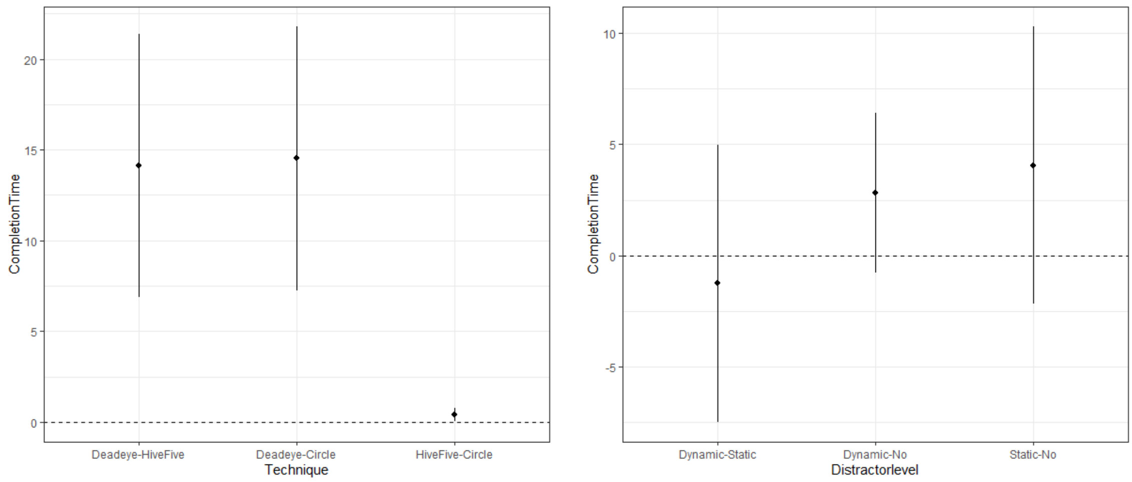
for Deadeye the static distractorlevel has the highest completion time. In addition, the no and the dynamic distractorlevel have a similar completion time whereas the static distractorlevel has the lowest for the Circle. However, the latter could be due to some hardware struggles with clicking on the right target. The former might be due to the chosen randomness of the distractorlevels.

For H1 and H2, we initially determined whether there are significant differences for the completion time among the techniques and distractorlevels. Based on our study design and the fact that the completion time data was not normally distributed, we decided for an aligned ranks transformation (ART) ANOVA [WFGH11]. This is a non-parametric approach for analysing main and interaction effects. Since the results (see Table 5.2) showed significant differences for both our independent variables (technique and distractorlevel), we did a pairwise comparison for both of them. For the

	F value	P value
Technique	75.986	< 2.22e-16
Distractorlevel	49.329	< 2.22e-16
Technique:Distractorlevel	13.967	3.1733e-10

Table 5.2: ART ANOVA [WFGH11] results for the completion time.

pairwise comparison, we plotted the confidence interval for the difference between the means (see Figure 5.2) for each technique combination (see Figure 5.2a) [CF05; Cum14]. Since none of the confidence intervals contain the zero, we can say that there is a difference in the mean completion time between each combination of the techniques. However, the distance between the confidence interval of the pairwise comparison of HiveFive and Circle to the zero is rather low compared to the distance of the confidence interval of Deadeye with both of the other techniques. Therefore, the effect of the difference in mean completion time of HiveFive and Circle might be quite low. In addition, we did a pairwise comparison of the distractorlevel by plotting the confidence interval for the difference between their means (see Figure 5.2b) [CF05; Cum14]. In this case, all confidence intervals contained the zero, which means that it is possible that there is no difference between the mean completion time of the distractorlevels.



(a) Pairwise comparison of techniques. **(b)** Pairwise comparison of distractorlevels.

Figure 5.2: Confidence interval for the difference of means [CF05; Cum14] for the completion time.

H1: The completion time increases for all techniques with increasing distractorlevel In H1 it is necessary to pairwise compare the distractorlevels (no vs. static vs. dynamic) for each technique. While all of the confidence intervals of the overall pairwise distractorlevel comparison suggest there might be no difference between them, the ART ANOVA [WFGH11] found an overall significant effect for distractorlevel. Additionally, it also found a significant interaction effect between the techniques and the distractorlevel. Due to the results of the ART ANOVA, we decided for a pairwise comparison of the distractorlevel for each technique. An additional comparison can be seen in Figure 5.1a, when comparing the plotted distractorlevel for each technique, especially for Deadeye and HiveFive.

We conducted a Wilcoxon signed-rank test [Wil45] with Bonferroni correction for the pairwise comparison of the distractorlevel as the data was not normally distributed. The results can be seen in Table 5.3. For Deadeye, it is only significant when comparing the no and dynamic distractorlevel. However, there is no significant difference between the static and dynamic distractorlevel. By looking at the confidence intervals of the mean in Figure 5.1a, it can be seen that the confidence

Deadeye			HiveFive			Circle		
	Dynamic	No		Dynamic	No		Dynamic	No
No	**0.0077**	-	No	**0.017**	-	No	0.128	-
Static	1.0000	0.5316	Static	**0.011**	1.000	Static	**0.001**	0.173

Table 5.3: P values of the pairwise comparisons for H1 using Wilcoxon signed-rank test [Wil45] using Bonferroni correction (with values in bold for $p < 0.05$).

interval for the static distractorlevel is rather high and comprises the confidence interval of the dynamic distractorlevel. Therefore, there are great differences between the participants and it would be reasonable to conduct further studies which examine this. According to the significant difference of no and dynamic distractorlevel and the increase of the mean completion time, it is still possible that the dynamic distractors influenced the perception of the highlighting. For example, when bees or birds were within the field of view and occluding the highlighted apple since some participants stated this in the comment section. In addition, it is also necessary to further examine the influence of static distractors in general as we could not find any significant difference.

For the Circle, there is only a significant difference for static and dynamic distractorlevel. By considering the Figure 5.1a, it can also be seen that the completion time for the dynamic distractorlevel is greater than the static distractorlevel while the no and static distractorlevel are rather comparable. Thus, it is possible that the dynamic scene elements (e.g., birds or bees) distracted the participants from the target resulting in a higher completion time. For example, when they were within the field of view. In addition, due to no significant difference, we can assume that it does not make a difference if there is only one tree or multiple trees within the scene. This suggests that it is possible that there is a difference between distractors being next to the target and distractors that possibly occlude the target which was also mentioned by participants. Furthermore, there is also no significant difference between no and dynamic distractorlevels. The completion time of the no distractorlevel is also higher than the completion time of the static distractorlevel (see Figure 5.1b). However, this result could be due to some hardware problems as some participants struggled with successfully clicking on the target leading to a slightly higher completion time.

In contrast to that there are two significant differences for HiveFive. The completion time differs significantly between no and dynamic as well as for static and dynamic distractorlevel, while being not significantly different between no and static distractorlevel. In this case, we can say that for HiveFive static elements do not distract the participant while dynamic distractors do. This can also be seen in the comments of the participants, as they mainly noted that especially the bees were disturbing.

All in all, we cannot confirm H1. Since there are some significant differences between certain distractorlevels for the techniques, it is possible that the chosen distractors have some influence on the completion time. However, we assumed that there might be a significant increase of the completion time with an increase of the distractorlevel. As this is not true for any technique, it is necessary to conduct further user studies for a more unambiguous statement regarding the hypothesis.

H2: The completion time increases more for HiveFive with dynamic distractors than with other techniques For H2 we compared the completion time of the dynamic distractorlevel for all techniques. Therefore, we also conducted a Wilcoxon signed-rank test [Wil45] with Bonferroni correction for the pairwise comparison of the techniques which can be seen in Table 5.4. In addition, we plotted in Figure 5.3 the confidence interval for the difference of the mean between each technique [CF05; Cum14]. Both show a significant difference between Deadeye and the other two techniques. Additionally, there is no significant difference between the Circle and HiveFive. Besides this, Figure 5.1b indicates that Deadeye has the highest completion time for the dynamic distractorlevel. Therefore, we cannot confirm H2 as the completion time does not increase more for HiveFive. However, it is higher than for Circle and also the results of the confidence interval

	Circle	Deadeye
Deadeye	**5.6e-09**	-
HiveFive	0.07	**1.7e-08**

Table 5.4: P values of the pairwise comparisons for H2 using Wilcoxon signed-rank test [Wil45] using Bonferroni correction (with values in bold for p < 0.05).

and the Wilcoxon signed-rank test [Wil45] suggest that it is only barely not significant. Thus, it might be possible that dynamic distractors affect dynamic techniques like HiveFive more than static techniques as the Circle. It might be reasonable to further investigate the influence of dynamic distractors on static and dynamic visual attention techniques. Additionally, it is worth studying Deadeye for further virtual environments as its completion time is rather high in general compared to both other techniques. While this could be due to the novelty of the technique or the missing visual feedback, there could be also other reasons. Even though there are significant differences between the completion time of Deadeye and both other techniques, we cannot be sure whether the difference results due to distractions by our chosen distractors or other reasons.

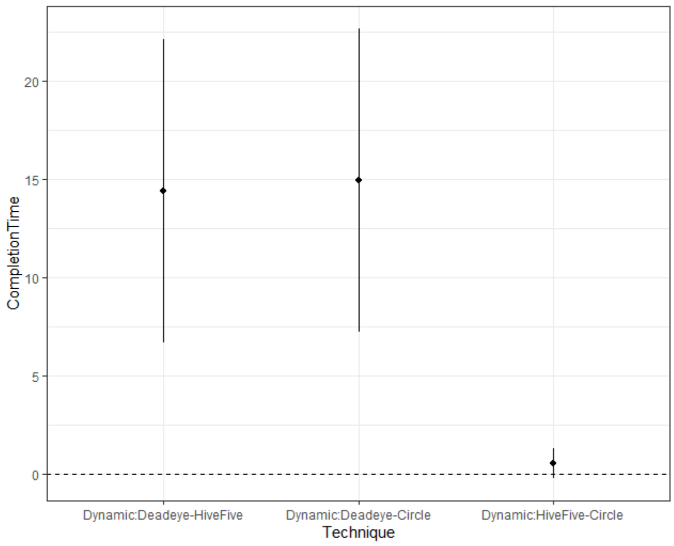


Figure 5.3: Confidence interval for the difference of means [CF05; Cum14] for completion time between techniques with dynamic distractorlevel.

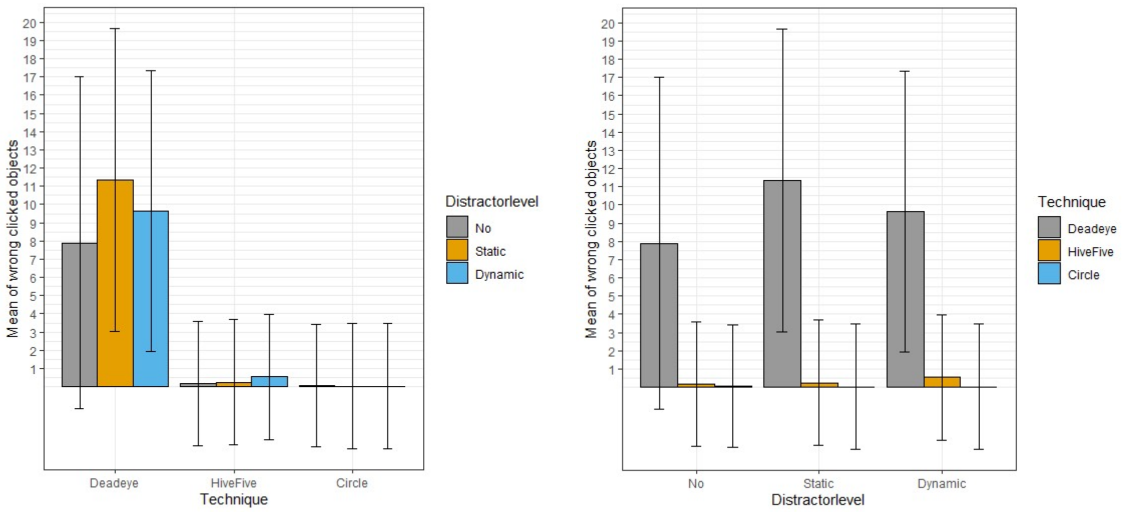
5.2 Wrong Clicked Objects

Besides the completion time of the targets, we also measured the wrong clicked objects per target of the participants. For this, we counted the number of clicked apples on the target tree (i.e., the tree that included the target) that the participants falsely perceived as a target. We again measured the wrong clicked objects per target resulting in 90 values per participant. Instead of calculating the mean, we added up all wrong clicked objects per distractorlevel for each target to aggregate the data to nine values per participant. For the evaluation, we calculated the mean and standard deviation for the distractorlevels and techniques (see Table 5.5).

Distractorlevel	Deadeye	HiveFive	Circle
No	M= 7.90 (SD= 24.44)	M= 0.16 (SD= 9.19)	M= 0.06 (SD= 9.03)
Static	M= 11.33 (SD= 22.29)	M= 0.23 (SD= 9.22)	M= 0.03 (SD= 9.19)
Dynamic	M= 9.63 (SD= 20.67)	M= 0.53 (SD= 9.15)	M= 0.03 (SD= 9.21)

Table 5.5: Mean and SD of wrong clicked objects by technique and distractorlevel.

By looking at the mean and confidence interval of the techniques and distractorlevel (see Figure 5.4), we can see a similar distribution as for the completion time (see Figure 5.1). As before, Deadeye has the highest amount of wrong clicked objects, with the highest peak for the static distractorlevel. In addition, it has a rather high confidence interval and thus, several variations between the participants. As expected, the amount of wrong clicked objects is rather low for Circle and HiveFive. For the Circle, there is a small increase for the no distractorlevel. However, we assume this might not be due to the distractorlevel but instead due to hardware struggles for less VR experienced participants. Interestingly, the mean increases with increasing distractorlevel for HiveFive. While the difference between the no and static distractorlevel is quite low, it is more noticeable between the static and dynamic distractorlevel. Hence, we can assume that dynamic distractors influence the perception of the target resulting in clicking on wrong objects. It is also possible that due to the size of the spheres of HiveFive, participants had struggles with aiming at the target. Thus, they clicked on the right target several times. Due to latency, the application recognizes a successful click, highlights the next target, but the participants still click on the previous target. While this is negligible for techniques with visual feedback like HiveFive and Circle, this has a greater negative impact on techniques like Deadeye resulting in higher numbers. It would also explain the differences (i.e.,



(a) Ordered by technique. **(b)** Ordered by distractorlevel.

Figure 5.4: Error bars showing the 95% confidence interval for the wrong clicked objects.

high confidence interval) between the participants as we also had differences in the VR experience. According to this, it might be necessary to further examine the influence of VR experience and its learning effects on the count of wrong clicked objects.

5.3 NASA TLX

Besides measuring the completion time and wrong clicked objects, the participants had to fill out a questionnaire containing the NASA TLX, some own questions as well as a text field for further remarks. Based on our study design, the participants had to fill out the questions nine times after completing each task. In this section, we present and discuss the results for the subscales of the NASA TLX and thus, the results of H3. Since the results of the mental demand and effort are relevant for H3, we first discuss them as one section with H3. This is followed by the remaining subscales as well as an overall task load subsection.

5.3.1 Mental Demand and Effort

Table 5.6 and Table 5.7 show respectively the mean and the standard deviation for mental demand and effort.

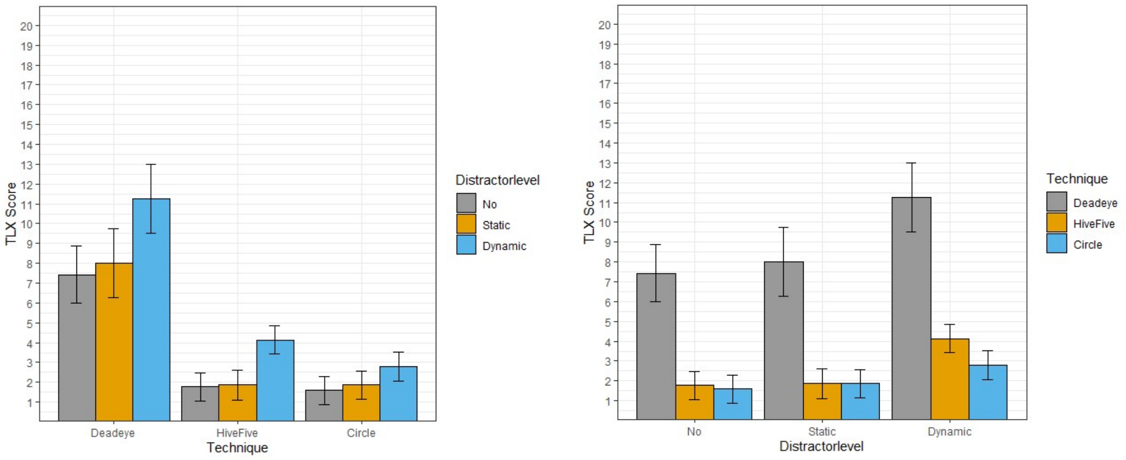
Mental Demand			
Distractorlevel	Deadeye	HiveFive	Circle
No	M= 7.43 (SD= 3.84)	M= 1.77 (SD= 1.86)	M= 1.60 (SD= 1.91)
Static	M= 8.00 (SD= 4.63)	M= 1.87 (SD= 2.03)	M= 1.87 (SD= 1.89)
Dynamic	M= 11.27 (SD= 4.63)	M= 4.13 (SD= 1.88)	M= 2.80 (SD= 1.95)

Table 5.6: Mean and SD of Mental Demand.

Effort			
Distractorlevel	Deadeye	HiveFive	Circle
No	M= 7.60 (SD= 3.73)	M= 3.00 (SD= 1.78)	M= 2.40 (SD= 2.36)
Static	M= 7.70 (SD= 4.24)	M= 2.57 (SD= 1.93)	M= 2.43 (SD= 2.21)
Dynamic	M= 10.70 (SD= 4.13)	M= 4.67 (SD= 1.69)	M= 3.67 (SD= 2.24)

Table 5.7: Mean and SD of Effort.

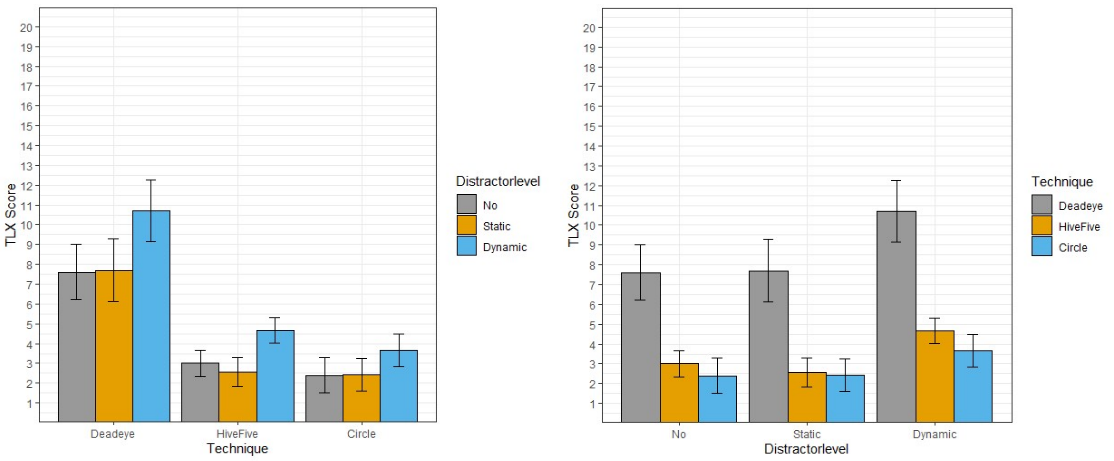
We also plotted the mean and the confidence interval for both subscales for technique and distractorlevel. For the mental demand (see Figure 5.5), Deadeye has again the highest values followed by HiveFive and finally the Circle (see Figure 5.5a). Furthermore, it increases for an increase in distractorlevel in every technique. While the increase is relatively small from no to static distractorlevel, it increases more for static to dynamic distractorlevel. Therefore, we can assume that the distractors have an influence on the mental state of the participants. It is possible that due to the movement of the distractors the participants had to do more mental task load to find the highlighted target, especially when the distractors are within the field of view. Since Deadeye is a rather subtle technique and due to the additional distractors it might be less obvious. Additionally, it is possible that the movement of the distractors is more demanding if the technique is also motion-based. This can also be seen in Figure 5.5b, since their difference for the dynamic distractorlevel increases while it is rather small for the no and static distractorlevel.



(a) Ordered by technique.

(b) Ordered by distractorlevel.

Figure 5.5: Error bars showing the 95% confidence interval for the mental demand (0: very low, 20: very high).



(a) Ordered by technique.

(b) Ordered by distractorlevel.

Figure 5.6: Error bars showing the 95% confidence interval for the effort (0: very low, 20: very high).

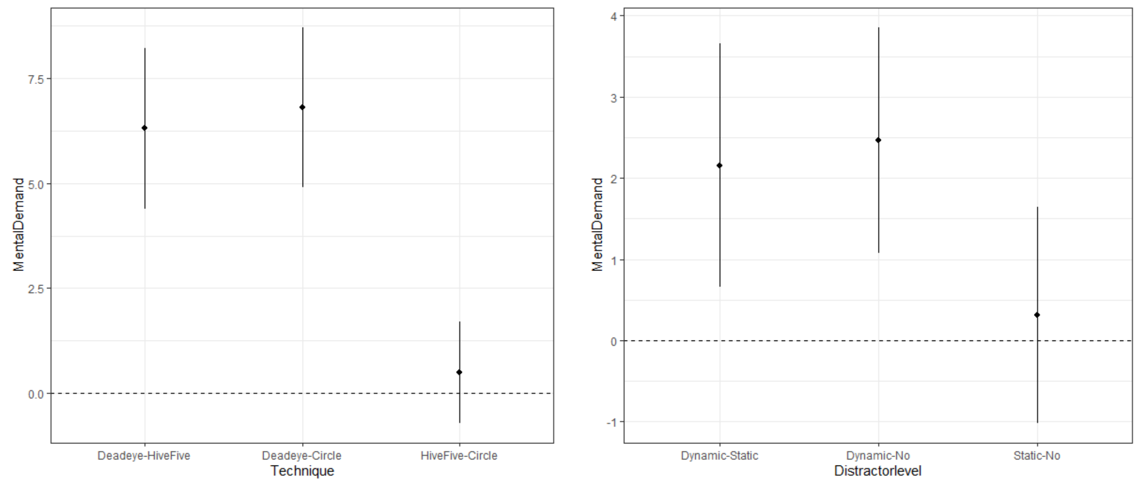
A similar distribution exists also for the effort (see Figure 5.6). Once again the values of HiveFive and Circle are rather comparable with a huge difference to Deadeye. Additionally, the values of no and static distractorlevels in each technique are comparable. Surprisingly, the static distractorlevel for HiveFive is smaller than the value for the no distractorlevel. However, this could be due to learning effects or the randomised order of the distractorlevel. It is also possible that the effort is smaller, because the participants did not perceive a difference between the distractorlevel and did not know their previous rating. Hence, they accidentally gave smaller or respectively, larger value. As before, there is a greater difference between the static and dynamic distractorlevel for each technique. Thus, the dynamic distractorlevel ist not only more demanding but the participants also had to work harder to find the highlighted target. Both subscales show that our distractorlevels do what they were intended to, since the participants perceived differences between them. However, the difference is generally greater for additional movement within the scene.

As for the completion time, we conducted an ART ANOVA [WFGH11] for the mental demand and effort to determine significant differences which can be seen in Table 5.8. We can see that for mental demand and effort, the techniques and distractorlevels have significant differences. In addition, both have a significant interaction effect with a greater value for effort.

Mental Demand			**Effort**		
	F value	P value		F value	P value
Technique	149.381	< 2.22e-16	Technique	120.8549	< 2.22e-16
Distractorlevel	28.031	1.248e-11	Distractorlevel	22.3764	1.2999e-09
Technique:Distractorlevel	4.258	0.0024101	Technique:Distractorlevel	3.0772	0.017022

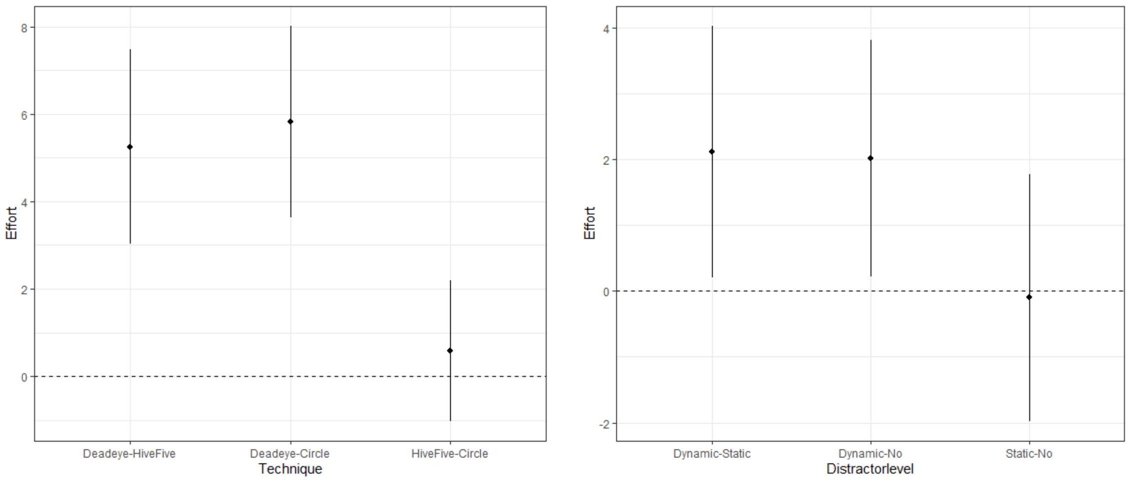
Table 5.8: ART ANOVA [WFGH11] results for the mental demand and effort.

Due to these results, we also did a pairwise comparison for both independent variables by plotting the confidence interval for the difference between the means for mental demand (see Figure 5.7) and for effort (see Figure 5.8) [CF05; Cum14]. For both, mental demand (see Figure 5.7a) and effort (see Figure 5.8a), we can see that there is a difference in the mean between Deadeye and both other techniques. We assume for both that the effect is rather high as the confidence interval has a greater distance to the zero (slightly higher for Circle than for HiveFive). However, it is possible that there is no difference in the mean between Circle and HiveFive as this comparison contains the zero. By considering the distractorlevel for mental demand (see Figure 5.7b) and effort (see Figure 5.8b), the following can be seen. There is a difference in the mean for the dynamic distractorlevel with each of



(a) Pairwise comparison of techniques. **(b)** Pairwise comparison of distractorlevels.

Figure 5.7: Confidence interval for the difference of means [CF05; Cum14] for mental demand.



(a) Pairwise comparison of techniques.

(b) Pairwise comparison of distractorlevels.

Figure 5.8: Confidence interval for the difference of means [CF05; Cum14] for effort.

both other distractorlevels (no, static) but that there is none between the no and static distractorlevel. Hence, we assume that in our case a difference between one and many trees (static distractors) does not influence the task load. The distance to the zero line of the comparison of dynamic and no or static distractorlevel is greater for mental demand than for effort. Therefore, we assume that the effect is also greater for mental demand than for effort. As expected, the distance to the zero line is greater for the difference between no and dynamic distractorlevel than for static and dynamic distractorlevel for the mental demand. This is reasonable, due to the increasing distractions within the scenes (no vs. static vs. dynamic). However, this is not the case for effort. In general, these results for the pairwise comparison of the distractorlevel (but also for the technique) are reasonable, since we saw similar trends in Figure 5.5 and Figure 5.6.

H3: Deadeye has higher mental demand and effort compared to other techniques with increasing distractorlevel We saw that both, Circle and HiveFive, have a significant difference to Deadeye for mental demand (see Figure 5.7a) and effort (see Figure 5.8a). Besides, we also saw that the mental demand (see Figure 5.5b) and effort (see Figure 5.6b) are generally greater for Deadeye than for HiveFive and Circle for each distractorlevel. Hence, we conducted a pairwise Wilcoxon signed-rank test [Wil45] using Bonferroni correction to determine whether these differences between the techniques for each distractorlevel are significant. Table 5.9 and Table 5.10 show their results.

For both, there are significant differences for each distractorlevel between Deadeye and HiveFive as well as between Deadeye and Circle. Thus, we can confirm H3. This confirms in our case that Deadeye is more demanding and needs more concentration. One reason for that can be that while Deadeye is rather subtle and does not have any visual feedback, both other techniques have visual feedback. It is also possible that Deadeye is perceived more demanding due to its novelty and stereoscopic restriction. Both other techniques use known features (motion, colour) to highlight the target. However, it might be reasonable to further examine reasons for the higher demand and also learning effects. It is possible that this high demand decreases when using Deadeye more often, since participants mentioned that they could perceive Deadeye better after a few targets.

No			Static			Dynamic		
	Deadeye	HiveFive		Deadeye	HiveFive		Deadeye	HiveFive
HiveFive	**3.6e-05**	-	HiveFive	**0.00019**	-	HiveFive	**0.00025**	-
Circle	**5.4e-05**	1	Circle	**0.00054**	1.00000	Circle	**1.6e-05**	0.25109

Table 5.9: P values of the pairwise comparisons of **mental demand** for H3 using Wilcoxon signed-rank test [Wil45] using Bonferroni correction (with values in bold for $p < 0.05$).

No			Static			Dynamic		
	Deadeye	HiveFive		Deadeye	HiveFive		Deadeye	HiveFive
HiveFive	**0.00281**	-	HiveFive	**0.0158**	-	HiveFive	**0.00099**	-
Circle	**0.00012**	1.00000	Circle	**0.0059**	1.0000	Circle	**8.9e-0.05**	0.87088

Table 5.10: P values of the pairwise comparisons of **effort** for H3 using Wilcoxon signed-rank test [Wil45] using Bonferroni correction (with values in bold for $p < 0.05$).

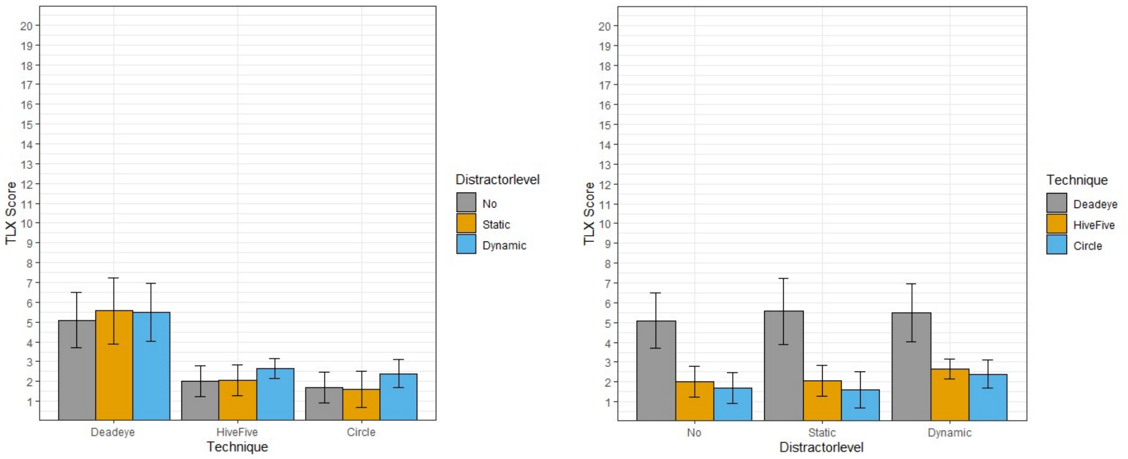
Interestingly, even though there is no significant difference between HiveFive and Circle in none of the distractorlevels, their p-value decreases with an increase in distractorlevel. This can also be seen in Figure 5.5b for mental demand and in Figure 5.6b for effort. Therefore, an increase in distractorlevel is generally more demanding for motion-based techniques than for static techniques. It is possible that similar features (e.g., moving spheres for HiveFive and flying bees) of the technique and distractors have an influence on the perception of the target highlighting when being in front of the target and thus, are more demanding. As our static distractors did not occlude any target or were in front of the technique, it did not need a higher demand to find the target. However, to confirm this effect it is necessary to conduct further studies.

5.3.2 Physical Demand

Besides the mental demand, the NASA TLX also includes a subscale for the physical demand. The mean and standard deviation can be seen Table 5.11. In general, the distribution of the physical demand per technique and per distractorlevel is comparable to the mental demand (see Figure 5.9). This means that Deadeye has higher values than both other techniques whereas the values for HiveFive and Circle are rather similar, with Circle being slightly smaller than HiveFive (see Figure 5.9b). A reason for that might be that Deadeye due to the stereoscopic effect of displaying the target to one eye and not to the other can be more physically demanding resulting in higher values. For Deadeye, the values of the physical demand are smaller compared to the mental demand. This is not really surprising as we assume that the task did not have any physically demanding interactions. In the case of Deadeye the difference between the no and the static distractorlevel is greater than the difference between the static and dynamic distractorlevel. Interestingly, the static distractorlevel is slightly greater than the dynamic distractorlevel (similar trend as for the completion time, see Figure 5.1a). However, for the physical demand this difference is only barely noticeable (see Figure 5.9a). This result could also be due to accidentally choosing one higher or smaller number and not due to a personal opinion. For HiveFive and Circle, as before for mental demand (see Figure 5.5a) and effort (see Figure 5.6a), the difference between no and static distractorlevel is rather small while the difference between the static and dynamic distractorlevel is greater (see Figure 5.9a). The difference of the no and static distractorlevel might be due to the same reason as for Deadeye between the static and dynamic distractorlevel. Therefore, we can conclude that the dynamic distractorlevel is not only mentally more demanding but also physically more demanding. Especially for HiveFive and Circle, it would be interesting to further investigate reasons for the rather high values for the physical demand.

Distractorlevel	Deadeye	HiveFive	Circle
No	M= 5.10 (SD= 3.75)	M= 2.03 (SD= 2.09)	M= 1.70 (SD= 2.12)
Static	M= 5.57 (SD= 4.45)	M= 2.07 (SD= 2.05)	M= 1.60 (SD= 2.43)
Dynamic	M= 5.50 (SD= 3.93)	M= 2.67 (SD= 1.38)	M= 2.40 (SD= 1.93)

Table 5.11: Mean and SD of the physical demand by technique and distractorlevel.



(a) Ordered by technique.

(b) Ordered by distractorlevel.

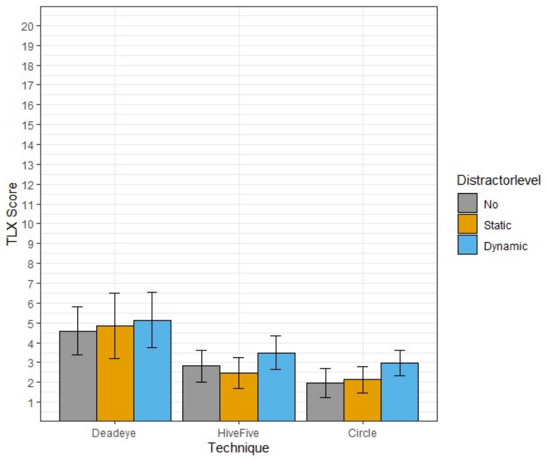
Figure 5.9: Error bars showing the 95% confidence interval for the physical demand (TLX Score ranging from 0 = very low to 20 = very high).

5.3.3 Temporal Demand

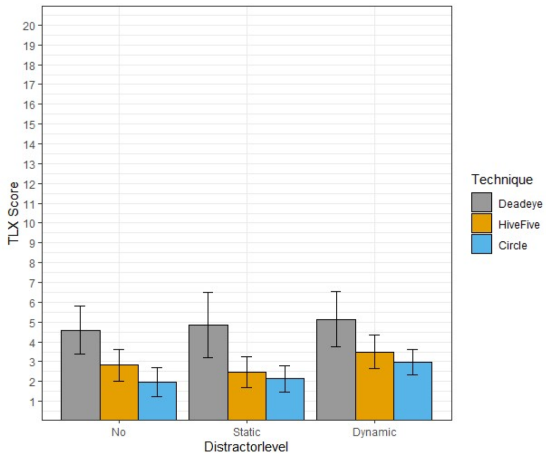
The values of the temporal demand (see Table 5.12, Figure 5.10) are quite similar to the value of the physical demand (see Table 5.11, Figure 5.9). Deadeye still has the greatest values followed by HiveFive and the Circle. As for the physical demand, the values of Deadeye are generally smaller and more comparable to the values of HiveFive and Circle as for the mental demand (see Figure 5.5) or the completion time (see Figure 5.1). While we expected the general distribution of temporal demand due to the distributions of the other values, the overall amount is still surprising. We did not include any timer or other time dependent variable visible for the participants and also did not tell the participants to hurry. We also did not tell them explicitly that we measured their completion time. In general, the values for Deadeye only slightly increase with an increase of the distractorlevel. As before, for Circle there is a small increase for no to static distractorlevel while it increases much more for static to dynamic distractorlevel. For HiveFive, the difference between the dynamic distractorlevel and both other distractorlevels is also increasing as for Circle. Surprisingly, the mean of the static distractorlevel is smaller than for the no distractorlevel. It could be that the participants felt more calm and relaxed when being in the static scenario since more scene elements (i.e., trees) made the scene more realistic. In contrast, only perceiving one tree might be more unrealistic as in nature several trees instead of one exist. However, we did not measure the feeling of presence and cannot examine this assumption. In case of the increase for dynamic distractorlevel, participants could be more stressed and feel time pressured due to quick and stressful movements of the distractors. Both assumptions need to be examined in another study.

Distractorlevel	Deadeye	HiveFive	Circle
No	M= 4.60 (SD= 3.27)	M= 2.83 (SD= 2.13)	M= 1.97 (SD= 1.94)
Static	M= 4.87 (SD= 4.39)	M= 2.47 (SD= 2.06)	M= 2.13 (SD= 1.79)
Dynamic	M= 5.13 (SD= 3.75)	M= 3.50 (SD= 2.23)	M= 2.97 (SD= 1.73)

Table 5.12: Mean and SD of the temporal demand by technique and distractorlevel.



(a) Ordered by technique.



(b) Ordered by distractorlevel.

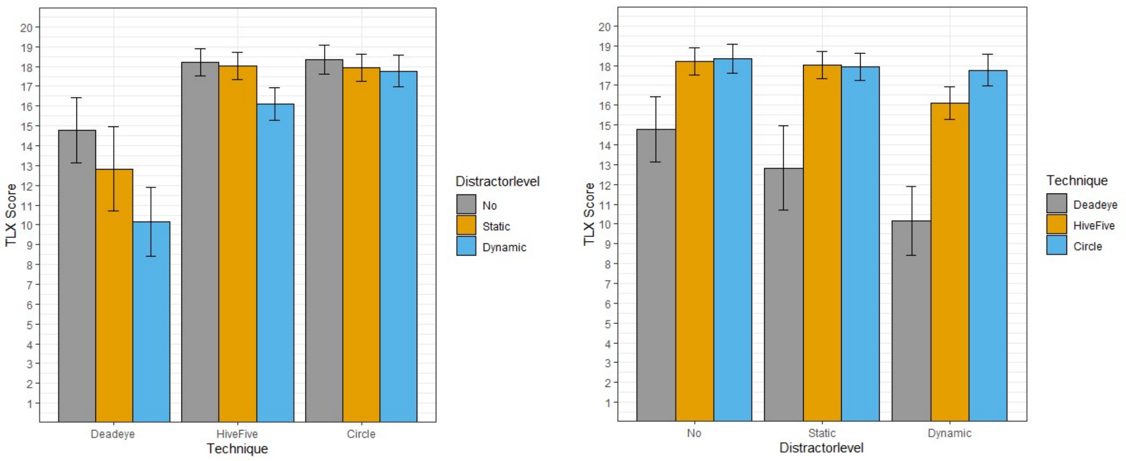
Figure 5.10: Error bars showing the 95% confidence interval for the temporal demand (TLX Score ranging from 0 = very low to 20 = very high).

5.3.4 Performance

The performance has a similar trend as the other subscales of the NASA TLX but due to reversed meaning of the range we are aiming for higher values (i.e., a perfect performance). The mean and the standard distribution can be seen in Table 5.13 and the distribution of mean with confidence interval is pictured in Figure 5.11. In general, the mean of the performance is higher for HiveFive and Circle than for Deadeye. Circle has rather comparable values with a small tendency to decrease for an increase of the distractorlevel. For HiveFive the mean is more decreasing for dynamic distractorlevel compared to no and static distractorlevel. Therefore, it is possible that dynamic distractors influence the task completion for HiveFive since there was also a significant difference of the distractorlevels for the completion time (see Table 5.3). One reason could be the motion-based features of both distractors and the technique. This is also supported by the values of the Circle as they are rather constant through the distractorlevels (see Figure 5.11a). However, it would be reasonable to examine the factors of the influence. The values of Deadeye show more clearly a rapid decrease with an increase of distractorlevels. Hence, it might be that the subtle behaviour of the technique leads to a decrease of confidence in one's own performance.

Distractorlevel	Deadeye	HiveFive	Circle
No	M= 14.77 (SD= 4.41)	M= 18.20 (SD= 1.87)	M= 18.33 (SD= 1.95)
Static	M= 12.83 (SD= 5.69)	M= 18.03 (SD= 1.79)	M= 17.93 (SD= 1.87)
Dynamic	M= 10.17 (SD= 4.69)	M= 16.10 (SD= 2.22)	M= 17.77 (SD= 2.18)

Table 5.13: Mean and SD of the performance by technique and distractorlevel.



(**a**) Ordered by technique.

(**b**) Ordered by distractorlevel.

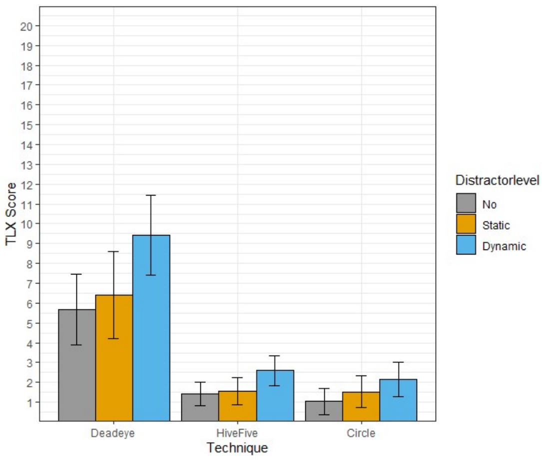
Figure 5.11: Error bars showing the 95% confidence interval for the performance (TLX Score ranging from 0 = failure to 20 = perfect).

5.3.5 Frustration

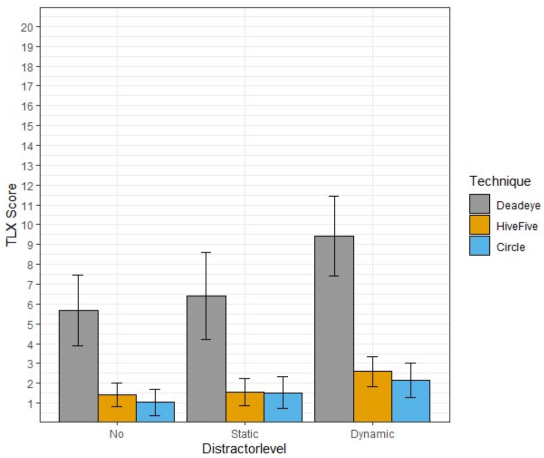
The last subscale of the NASA TLX questionnaire is the frustration and its mean and standard deviation can be seen in Table 5.14. Its distribution regarding the techniques and distractorlevel can be viewed in Figure 5.12. As prior, Deadeye has greater values than HiveFive and Circle. In addition, the values of HiveFive and Circle differ mostly between the no and dynamic distractorlevel (see Figure 5.12b). For Circle, the increase of the mean is rather similar for no and static as well as for static and dynamic distractorlevel. Thus, it is reasonable that more scene elements can increase the tendency of a negative mental state but for the Circle, this state is still rather small (even for the dynamic distractorlevel, see Figure 5.12a). In contrast to that are the differences for Deadeye and for HiveFive. The addition of static distractors does not influence the perceived frustration as much as the addition of dynamic distractors in the scene (see Figure 5.12a). For both techniques, it is possible that distractors within the field of view and thus, potentially occluding the target influence the perception of the participant. It might be that participants therefore feel more stressed and annoyed. Once again it requires more user studies to confirm this assumption.

Distractorlevel	Deadeye	HiveFive	Circle
No	M= 5.67 (SD= 4.76)	M= 1.43 (SD= 1.60)	M= 1.03 (SD= 1.79)
Static	M= 6.40 (SD= 5.88)	M= 1.57 (SD= 1.85)	M= 1.53 (SD= 2.16)
Dynamic	M= 9.43 (SD= 5.39)	M= 2.60 (SD= 2.03)	M= 2.17 (SD= 2.33)

Table 5.14: Mean and SD of the frustration by technique and distractorlevel.



(a) Ordered by technique.



(b) Ordered by distractorlevel.

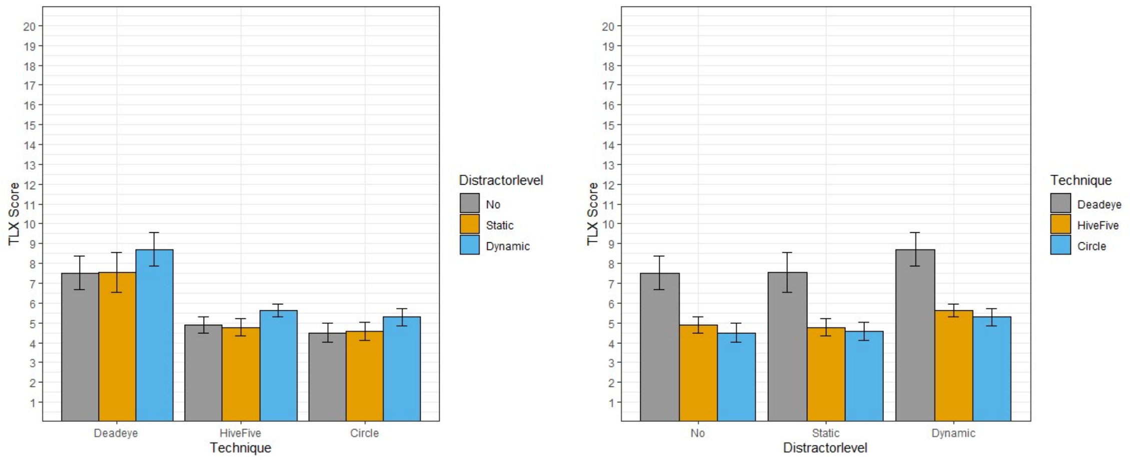
Figure 5.12: Error bars showing the 95% confidence interval for the frustration (TLX Score ranging from 0 = very low to 20 = very high).

5.3.6 Overall Taskload

Besides the specified values by the participants for the subscales of the NASA TLX, we also calculated the overall task load of the participants. Therefore, we calculated the mean of values for the given subscales (e.g., mental or physical demand) per participant and created another column per participant. For this data, we could again calculate the descriptive statistics which can be seen in Table 5.15. Additionally, we plotted the distribution with the confidence interval that can be viewed in Figure 5.13. Due to the overall similar distributions between the subscales, the given distribution for overall task load is not surprising. As before, the mean of Deadeye is generally higher as for HiveFive and Circle (see Figure 5.13b). It might be possible that Deadeye is in general more demanding than the other techniques, for example, due to the absent visual feedback. However, it could also be the case that this higher demand is due to the novelty and unknown manner of the technique. It is reasonable to further investigate its learning effects or the demand of Deadeye over a longer period. Additionally, we did not tell the participants in advance which kind of highlighting they could expect, thus it might also be interesting to study the differences when participants know how the target will be highlighted. While the difference between no and static distractorlevel is for all techniques rather similar, it increases between static and dynamic distractors (see Figure 5.13a). As this result was and is rather similar for the most subscales, it is reasonable to study which distractors influenced the given answers. Since we included several distractors, a more finer increase of distractorlevels could be useful. However, as there are only few contributions that consider distractors, we wanted to generally differentiate between static and dynamic (or moving) scene elements. For this study it is possible that generally the dynamic distractors affected the task load of the participants in some way. Further studies could therefore try to examine which and how these distractors influenced it.

Distractorlevel	Deadeye	HiveFive	Circle
No	M= 7.53 (SD= 2.26)	M= 4.88 (SD= 1.09)	M= 4.51 (SD= 1.29)
Static	M= 7.56 (SD= 2.68)	M= 4.76 (SD= 1.17)	M= 4.58 (SD= 1.18)
Dynamic	M= 8.70 (SD= 2.26)	M= 5.61 (SD= 0.86)	M= 5.29 (SD= 1.15)

Table 5.15: Mean and SD of the overall taskload by technique and distractorlevel.



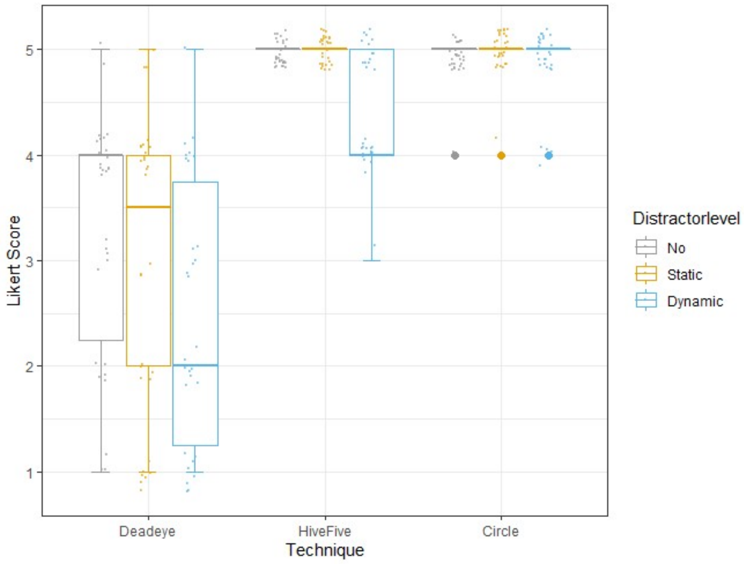
(a) Ordered by technique. **(b)** Ordered by distractorlevel.

Figure 5.13: Error bars showing the 95% confidence interval for the overall taskload (TLX Score ranging from 0 = very low to 20 = very high).

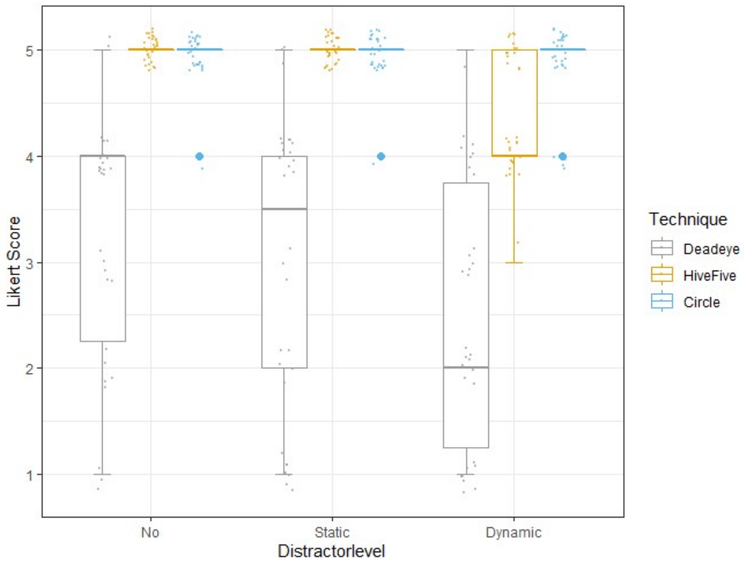
5.4 Own Questions

In addition, we integrated four questions with a 5-Likert Scale (1 = strongly disagree to 5 = strongly agree). We plotted the results as boxplots (see Figure 5.14 to Figure 5.17) and ordered them as before by techniques and distractorlevels. We also added some visual jitter to our boxplots to get further insights of the underlying distribution.

Q1 The first question states **'I could perceive highlighted object well'** and its results are presented in Figure 5.14. In general, the distribution is quite similar to the distribution of the performance (see Figure 5.11). While for Deadeye the minimum and maximum values of the boxplot use the whole range, it is quite constant for HiveFive and Circle (see Figure 5.14). One exception is the dynamic distractorlevel for HiveFive as the minimum values and interquartile range decrease and contain values up to 3 (see Figure 5.14b). The remaining distractorlevel of HiveFive and all distractorlevel for Circle keep their median during all conditions (MD = 5) with only some outliers for Circle (see



(a) Ordered by technique.

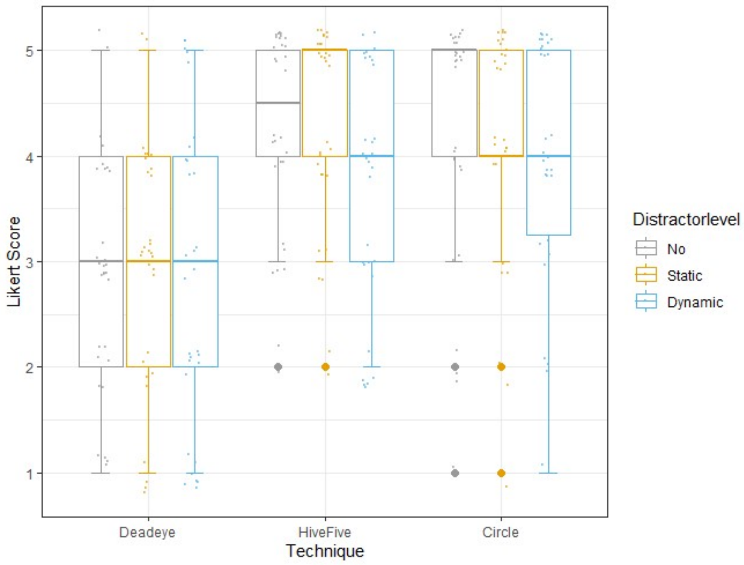


(b) Ordered by distractorlevel.

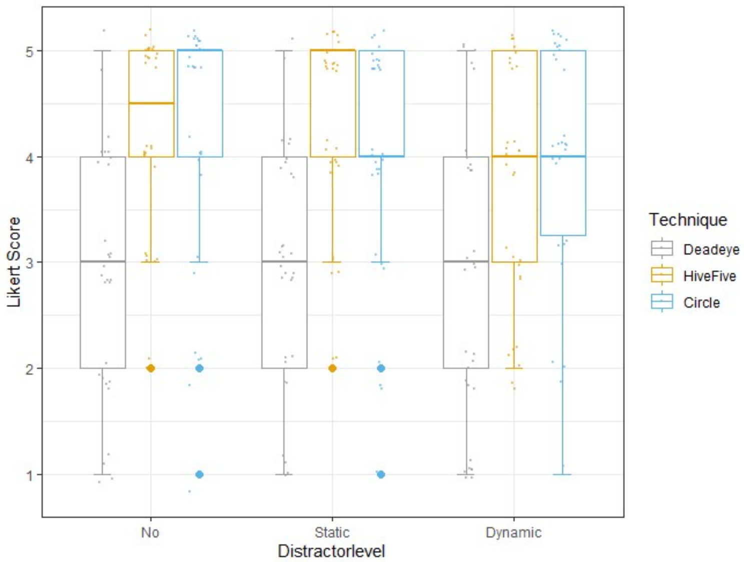
Figure 5.14: Boxplot with jittered points (as data distrubtion, small dots) and outliers (bigger dots) for 'I could perceive highlighted object well'.

Figure 5.14b). Thus, it could be that due to a similar behaviour of HiveFive and the distractors, the participants were hindered to perceive the target. This might be especially true for the bees and birds as their general moving behaviour acted as the swarm of HiveFive using the same algorithm (flocking algorithm by Reynolds [Rey87]). For an increase of the distractorlevel with Deadeye, the interquartile range and also the median decreases (MD = 4 to MD = 2, see Figure 5.14a). Therefore, we can suggest that dynamic distractors within the field of view make it more difficult to see the highlighting technique due to its subtle behaviour. Generally, we can say that Deadeye is less perceivable than HiveFive and Circle which might be due to its missing visual feedback within the scene or its novelty. However, it could be also due to some other factors. Additionally, we assume that the static and non-motion behaviour of the Circle is beneficial for the perception regardless which distractors are present.

Q2 The results of the second question **'Technique fits well into overall scene setting'** can be viewed in Figure 5.15. In contrast to the results for HiveFive and Circle, the boxplots of Deadeye remain constant. As for the first question (see Figure 5.14), they span within the entire range.



(a) Ordered by technique.

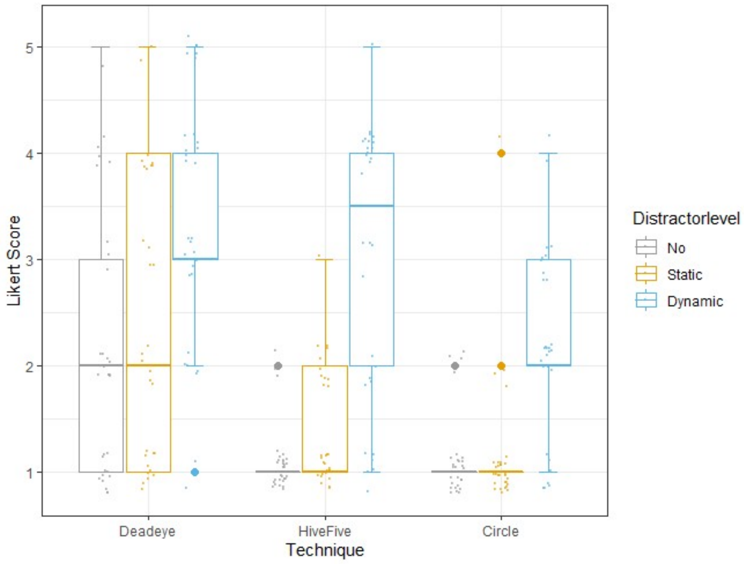


(b) Ordered by distractorlevel.

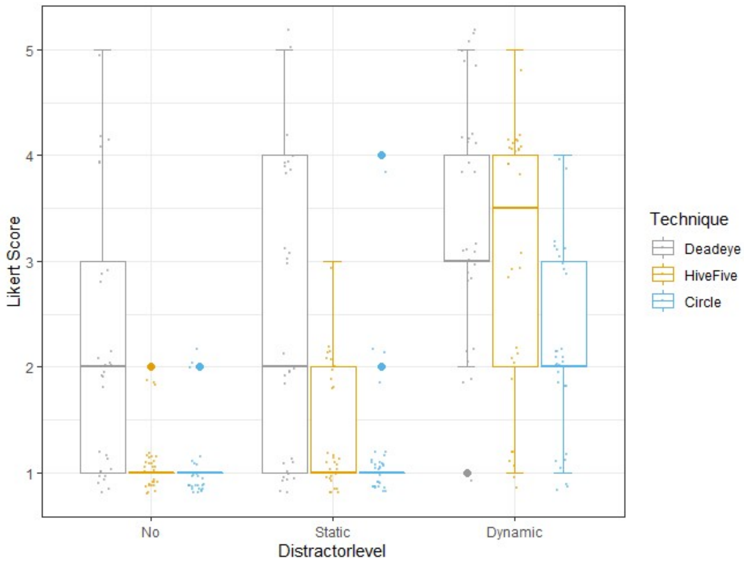
Figure 5.15: Boxplot with jittered points (as data distrubtion, small dots) and outliers (bigger dots) for 'Technique fits well into overall scene setting'.

However, unlike before the median and the interquartile range remain the same for all distractorlevels (MD = 3, see Figure 5.15a). Thus, it is possible that participants could not decide whether the technique fits or fits not within the environment. It could also be that the type of highlighting is difficult to assess in terms of suitability due to its unique visual properties. Furthermore, the results are quite high for HiveFive and Circle (median between 4 and 5, see Figure 5.15a). In terms of HiveFive, the distribution of most values remains constant for the no and static distractorlevel. For the dynamic distractorlevel the interquartile range includes decreases to lower values and also the minimum value decreases up to 2 (see Figure 5.15b). Interestingly, although the interquartile range does not change between the no and static distractorlevel, the median increases for a static distractorlevel (MD = 4.5 for no and MD = 5 for static distractorlevel, see Figure 5.15a). However, it decreases again for the dynamic distractorlevel (MD = 4, see Figure 5.15a). It is possible that an increase of natural scene elements (in our case static distractors e.g., trees, flower beds) may enhance the feeling of presence and realism for the participants. Thus, the increase of realism of the scene potentially gives the flock behaviour of HiveFive a more natural look. But the addition of moving scene elements (dynamic distractors) decreases this feeling again due to the more basic appearance of the HiveFive (yellow spheres) and hence, the flock does not fit as good as before any more. It would be interesting to examine whether the values change for the dynamic distractorlevel when using actual bees or ladybugs instead of spheres. However, it is necessary to conduct further studies for the actual reasons of these findings. The most interesting results are the values for Circle, as we expected a yellow circle to not fit in such a virtual environment. However, the values of all three distractorlevel are very high (see Figure 5.15a). For the no distractorlevel the median is also greater than the median of HiveFive. It decreases for the static distractorlevel. Both interquartile ranges are similar to HiveFive but have more outliers (see Figure 5.15b). The dynamic distractorlevel of Circle has the same median (MD = 4, see Figure 5.15b) as the one of HiveFive. However, while the interquartile range for Circle is generally smaller containing only higher values than HiveFive, its minimum value is definitely smaller (min = 3 (HiveFive), min = 1 (Circle)). It is reasonable to investigate the reasons for these values for such two-dimensional techniques in further studies. It might be also useful to study whether it is really necessary to have a more adapted technique to the virtual environment when techniques like Circle are perceived as well fitting.

Q3 The first two questions were rather general for the techniques and distractorlevels, the last two questions consider the scene elements within the different distractorlevel (i.e., especially the trees for the static scene and the bees, birds or the wind in the dynamic scene). For the first of the last two questions **'The current visual scene elements distracted me from the target search'** the results can be seen in Figure 5.16. For HiveFive and Circle the median as well as the minimum and maximum values remain for the no distractorlevel at one with some outliers at 2 (i.e., participants did mostly strongly disagree, see Figure 5.16b). Surprisingly, this is not the case for Deadeye. Even for the no distractorlevel, the range of the values use the whole scale although most values lie between 1 up to 3 with a median of 2 (see Figure 5.16b). Hence, we can assume that due to the subtle behaviour of Deadeye other potential targets may be distracting from finding the correct target. For a static distractorlevel of Deadeye the median remains but the interquartile range includes greater values up to 4. Furthermore, the interquartile range is higher and also the median is higher for the dynamic distractorlevel, but the whole range does not contain the whole scale but only contains values between 5 to 2 (see Figure 5.16a). Therefore, we see a trend that an increase in distractorlevel can distract participants more for the target search. We can also assume that our distractors fulfilled their task to distract participants. In terms of Deadeye it might be due to the subtle properties but



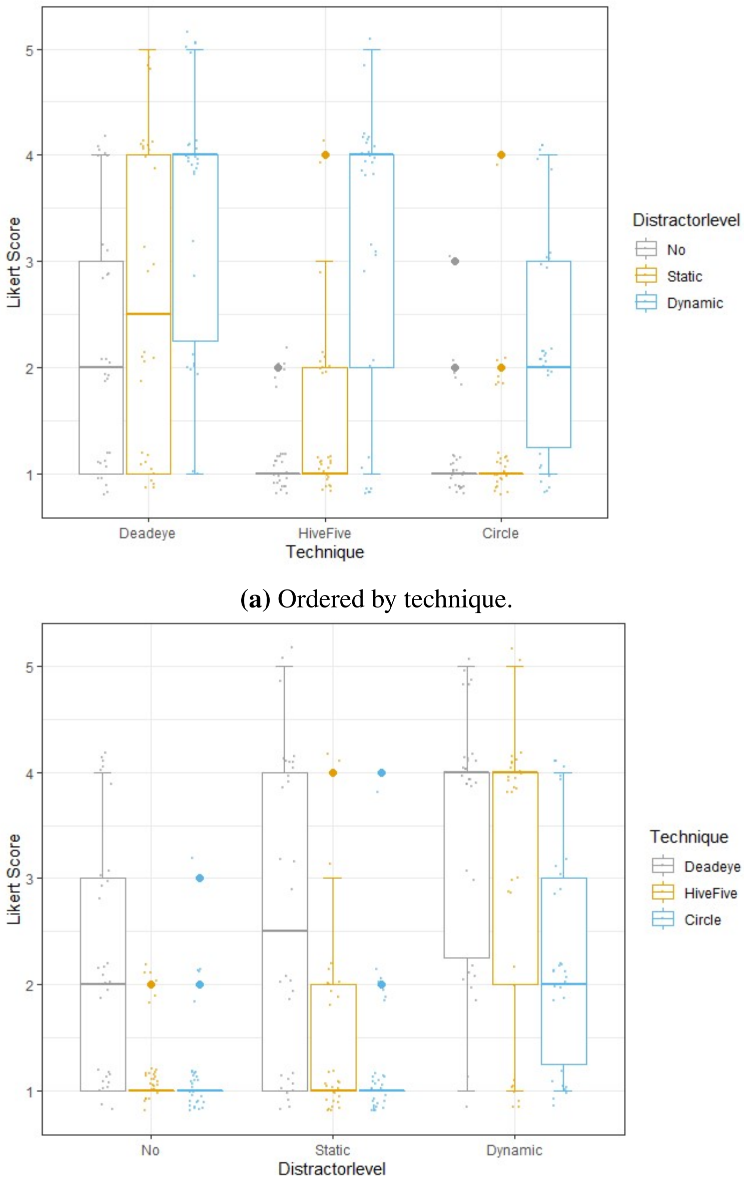
(a) Ordered by technique.



(b) Ordered by distractorlevel.

Figure 5.16: Boxplot with jittered points (as data distrubtion, small dots) and outliers (bigger dots) for ‘The current visual scene elements distracted me from the target search’.

could also include several other factors. A similar trend is also observable for HiveFive. Although the median remains at 1 for the static distractorlevel, the interquartile range increases to a value of 3 (see Figure 5.16b). For the dynamic distractorlevel also the median increases (MD = 3.5) and the minimum and maximum values use the whole range. In contrast to Deadeye, the interquartile range also includes lower values (see Figure 5.16a). In this case, we suggest that our distractors also distracted the participants for HiveFive, especially when the distractors had similar behaviours as the technique itself. Lastly, the static distractorlevel did not influence the target search for the Circle as its result remains similar to the no distractorlevel. Also for the static Circle, the values increase with presence of dynamic distractors (see Figure 5.16a). Even though the median only increases by one (MD = 1, see Figure 5.16b), the interquartile range and also the minimum to maximum value contain higher values. Hence, we assume that moving distractors are generally more distracting than static scene elements as they direct the attention less. However, it might be useful to also examine which distractors distract the most and which distractors are not perceived at all. In this case, it is reasonable to increase the distractors in smaller steps instead of just three large steps.



(a) Ordered by technique.

(b) Ordered by distractorlevel.

Figure 5.17: Boxplot with jittered points (as data distrubtion, small dots) and outliers (bigger dots) for 'The current visual scene elements made the task more difficult'.

Q4 Lastly, we did not solely ask whether the scene elements distracted but also whether **'The current visual scene elements made the task more difficult'**. We assumed that it is possible that even though they distracted the participant, they did not hinder the participants to complete the task. In general, the values for the last question in Figure 5.17 for all techniques are similar to the values of the question before as seen in Figure 5.16. There is only an increase of the median (MD = 4, see Figure 5.17a) for the dynamic distractorlevel of HiveFive and a few outliers more for Circle when there are no distractors which does not change the boxplot distribution. Therefore, we assume that an increase of distractorlevel makes the task more difficult (especially for Deadeye and HiveFive). This may be due to their novelty or the general properties as the common technique Circle did not have such an increase of values. However, it might be also due to other factors. Since there is not much research of distractors within VR up to this point, it might be reasonable to conduct more studies to give further insights. As mentioned before, it could be useful to find out which type of distractors distracts the most and also the reasons for that. Therefore, it is possible to use techniques within different scenarios due to their potential distraction resistance.

5.5 Qualitative Results

Besides the measurements of the completion time and wrong clicked objects, as well as the data from NASA TLX and our own questions, we gathered some qualitative data. First, we present the results of the further remarks text field, which was optional after each condition. Afterwards we consider the perceived mental and physical strains and the ranking of techniques of the post study questionnaire. For the evaluation of this data, we sorted the given texts by techniques and then, by similar topics. In addition, we deleted multiple occurrences of a similar answer and re-formulated it for a better understanding.

5.5.1 General Feedback

The optional text field could be used to give further remarks to a technique, reasons for some decisions or for further feedback. There were some general statements about the study design. For example, it was mentioned that a participant was quite surprised by the moving objects at first. They further explained that these objects made the task more comfortable and fitted well in the scene. A few participants said that they could focus easier on the task and ignore the distractors after several repetitions. While there were only few remarks and feedback about the general study design, there were many statements about the techniques in general.

In the case of Deadeye some participants stated that they did not like the technique and said that it was uncomfortable for the eyes. For the latter, they added that they could not imagine to use the technique in a long run (e.g., within a video game). One participant described it as *'scary'* while another reasoned that the scene felt more natural than with the Circle. Most participants described that Deadeye appeared to them as a glitch or a hardware, visual bug and thus, ignored it at first. Some participants also blamed their own eyes if an apple appeared more blurred. As Deadeye was also the first technique for some participants, one stated that they expected a more obvious highlighting.

Interestingly, several participants mentioned that the first targets were hard to find and they needed time to understand the highlighting. After some targets, they got an idea or claimed that they focused on the change between to targets as they perceived it to be more obvious than the highlighting itself. Therefore, we decided it might be reasonable to consider the completion time for the different rounds regardless of the distractorlevel. In Figure 5.18, we can definitely see that completion time strongly decreases between the first and the second run. In addition, it also decreases for the second and third run but not as much as for the first two runs. For the third run (see Figure 5.18), the completion time is quite comparable to the mean of completion time for HiveFive and Circle for certain distractorlevel (see Figure 5.1). Hence, it would be reasonable to study the learning effect but also the long-term usage of Deadeye. However, as our aim is to investigate the influence of distractors, we do not investigate this effect any further.

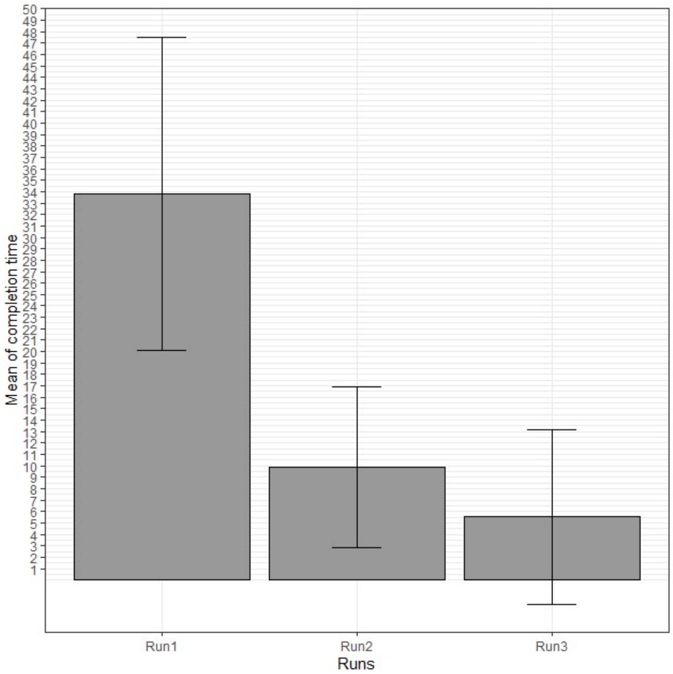


Figure 5.18: Mean of completion time per run of Deadeye.

As previously discussed in terms of the wrong clicked objects (see Section 5.2), some participants also remarked that it was difficult to be sure whether they successfully clicked on a right target as there was no visible feedback within the scene and thus, clicked a few times too often. In contrast to Deadeye, several participants said that HiveFive was easier due to the visual feedback or that they did not need any practice. However, some other participants also claimed that it was complicated to understand at first but after a few rounds their frustration decreased and it was easier to identify. For HiveFive, they mostly liked the combination of motion and its colour, as they argued that highlighting by movement was the easiest to perceive or the chosen colour provided a great contrast to the scene. Therefore, it might be reasonable to investigate the impact of colour on the perception of HiveFive as the chosen colour by Lange et al. [LSGB20] was very bright. Since Lange et al. [LSGB20] claim that they wanted to have a more immersive swarm adapted to the environment, it would be necessary to study how well the technique performs if it is illustrated as a real swarm of bees or lady bugs. Also the motion seemed unfavourable as some participants said it is more complicate to aim at the target and the smaller target space increased their frustration. Like Deadeye, some participants perceived that it was more difficult when the targets were located in the periphery. Nevertheless, some participants countered that HiveFive was still clearly perceptible even at greater distances. Besides HiveFive, several participants liked the Circle due to its good highlighting of the target. Contrary to Deadeye, some mentioned that the Circle was clearly and easily visible and as HiveFive did not need any practice. Additionally, they described it as nice and calm, unobtrusive and pleasant. Some participants also added that the bright colour appeared natural and eye-catching and fitted in the scene. Other participants contradicted or mentioned that its shape felt unnatural.

By comparing the Circle to HiveFive, participants described the Circle as less stressful or more boring. As previously seen in Figure 5.15, there were very mixed opinions about the fitting of the Circle into the scene. Some participants argued that the Circle fits in the scene as it is clearly perceivable with distractors or that the Circle generally fits more than HiveFive. This is especially of interest as we expected the Circle to be less immersive than HiveFive and thus, less suitable (or fitting) in this more natural and realistic virtual environment. However, we assume that our definition of suitable might be different to some participants. However, several other participants claimed that more subtle methods would be more suitable (for example, *'the Circle fits worse than Deadeye'*). Some also found it less immersive or obvious than other presented techniques but also said that it would highlight the target better or had a similar performance. One benefit to HiveFive was for some participants that they could easier aim at the target, because it surrounds the target without occluding it.

Additionally, some participants described that the difficulty of finding the target when highlighted by Deadeye varied between different targets. Some claimed that the perceived intensity of the highlighting varied with head turns or that targets in the periphery were harder to see as targets in near distances. Interestingly, one participant with a dominant left eye mentioned that right targets where generally more difficult to perceive than left targets and thus, performed head turns to perceive the targets. We also got the impression while conducting the study that participants with a dominant left eye needed more hints for Deadeye. While Krekhov and Krüger [KK19] claimed that there would be no difference of eye dominance, it might be reasonable to conduct further studies with more participants as they did not have that many participants, especially with a dominant left eye.

Some participants assessed also Deadeye regarding distractorlevels. In general, they perceived Deadeye to be more easy for only one tree as the target could only be on this tree. This was especially the case if Deadeye was the first technique, as participants were unsure whether the target could also be on the other trees. In addition, due to counterbalancing and randomness of the distractorlevel order, one participant falsely assumed the flower bed with polls as the target. They explained that the polls were similar to a common highlighting technique in video games. Besides the polls, some participants claimed that the dynamic objects, especially the bees, made the perception of the target more difficult as it was not always visible. However, some other participants contradicted this.

In case of HiveFive, several participants argued that the technique was the easiest to perceive in scenes when being the only movement within the scene. Furthermore, they did not perceive any difference for one or many trees. Additionally, some participants claimed that it would fit more in the environment with many trees or in the dynamic scene as it felt more unnatural and prominent in the *'empty'* (no distractorlevel) scene. This trend can also be seen when considering the boxplots for the third and the last questions of our own questions (see Figure 5.16 and Figure 5.17). For some participants the spheres appeared rather unnatural compared to the scene appearance of the static and dynamic distractorlevel and would like the technique if it would not consist out of *'big yellow* *balls'*. Interestingly, one participant claimed that HiveFive would directly guide the attention and therefore, *'felt more in a game rather than being in the real world'*. In terms of HiveFive, several participants claimed that the dynamic distractors, for example the bees or pollen, made the task more complicated. Especially the bees due to their similar appearance and colour were difficult to distinct from the technique.

While there were several mixed opinions for HiveFive and Deadeye, participants agreed that the distractors disturbed less or not at all for Circle. They argued that it could be better distinguished from the bees even though it had the same colour because of its static behaviour as a contrast to the motion of the bees. Someone also claimed that the bees still disturb but the Circle remains easy to recognize. However, a few participants said that the Circle is overlook-able or harder to perceive or more difficult to see when being occluded by leaves (i.e., due to the wind).

In general, it is really interesting to see that some dynamic distractors were mentioned quite often (e.g., wind, bees or pollen) while others were not (e.g., birds, the fallen apple). Therefore, we argue that it might be reasonable to further study the influence of distractors especially in their intensity. It is possible that there are several differences which could be important to know for potential applications as VR movies.

5.5.2 Strains

Since we were not sure about the potential impact of the VR application and the techniques, we asked the participants whether they were feeling any strains after the study. We also advised them to first go for a walk before driving home. We also assumed it might be possible that Deadeye could have a higher impact on the sight of participants due to the uncommon stereoscopic vision (e.g., seeing an object only on one eye). We added a 5-Likert Scale for physical (e.g., headache or eye pain) and mental strains (e.g., feeling dizzy) with the results shown in Figure 5.19. It can be seen

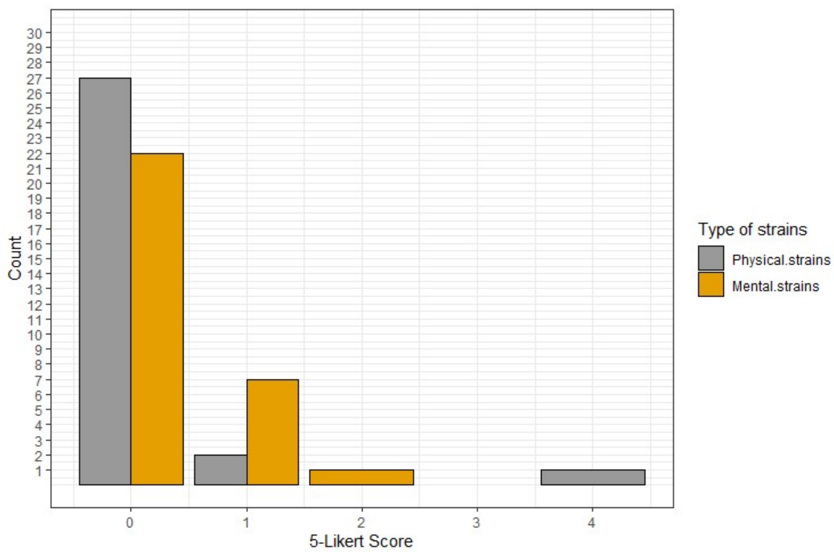


Figure 5.19: Perceived strains by participants.

that most participants did not have any strains at all. Whenever participants felt strains, they mostly perceived it as mental strains (as there are some higher values for mental strains, see Figure 5.19). Only one participant felt strong physical strains.

5.5.3 Ranking

Besides potential strains, we were also interested in the ranking of the techniques and reasons why the participant chose a certain rank for a technique. The distribution of ranks by techniques are shown in Figure 5.20. As most participants did either chose HiveFive or Circle as rank 1 or rank 2 it is not surprising that Deadeye is mostly on the last rank. They mostly described Deadeye when ranking them last as generally difficult or non-intuitive. Some participants also mentioned that it is more difficult to detect compared to the other techniques especially for the first targets or if being occluded by leaves or other objects. They also mentioned that it may be hard to find when the target is not within the field of view but in periphery. In contrast, it would also take longer to understand. Interestingly, some participants mentioned that it is discomforting for the eyes but as seen before (see Figure 5.19), they did not mention any physical eye pain. However, it is possible that the effect of discomfort does not have a long-term effect. One participant mentioned as a reason that Deadeye *'makes you go crazy, question your sanity first, then you suspect a graphical glitch and finally you [c]lick just to make it go away'*. Also other participants described Deadeye as creepy, frustrating or *'like a horror scene in which things behave strange'*. Some also called it a visual or technical bug.

In contrast, most people liked that HiveFive and Circle are easy to identify and did not need further instructions to understand how the target was highlighted. According to the participants, both techniques were also visible in the periphery. Generally, more participants ranked HiveFive on rank 1 than Circle, even though the difference is rather small (see Figure 5.20). Therefore, they reasoned that HiveFive was *'fun to work with'* and very clear due to its motion and colour. Some participants preferred HiveFive over Circle due to its satisfying and visually pleasing appearance while Circle was being rather unnatural and not visually appealing. In addition, they also suggested that it may be a great trade-off between guiding the attention and being adapted to the scene. In contrast, Circle was decreasing the immersion or harder to perceive with occlusions or moving objects.

Since the difference of HiveFive and Circle for rank 1 is rather smaller, there are also several reasons why people preferred Circle over HiveFive. While some people preferred motion over static behaviour, some argued that HiveFive would occlude the target object. Therefore, it would be harder to aim at the target whereas the Circle would frame the apple well. Additionally, some participants mentioned that the motion of HiveFive would be intrusive and more distracting. Other reasons were that the participant perceived the highlighting as a part of the scene and some mentioned that the task was harder when *'real bees'* were present. In contrast to that, one participant mentioned that HiveFive due its *'simple shapes help to keep grounded even though it was animated'*.

However, there are also a few people who ranked Deadeye first. Interestingly, the reasons for ranking Deadeye first or last barely differ. These participants liked that it was the most challenging or difficult and therefore, most interesting. Additionally they mentioned that Deadeye did not change the environment and that it was very good to perceive after some practice. One participant called it a *'great visual experience'*. A similar trend is also noticeable for Circle. While some people liked the easy and straightforward behaviour of Circle, some reason that it is easy and functional and ranked it last. Reasons for ranking HiveFive last were that the participants found the motion or the technique as a whole annoying especially in the dynamic scene with bees.

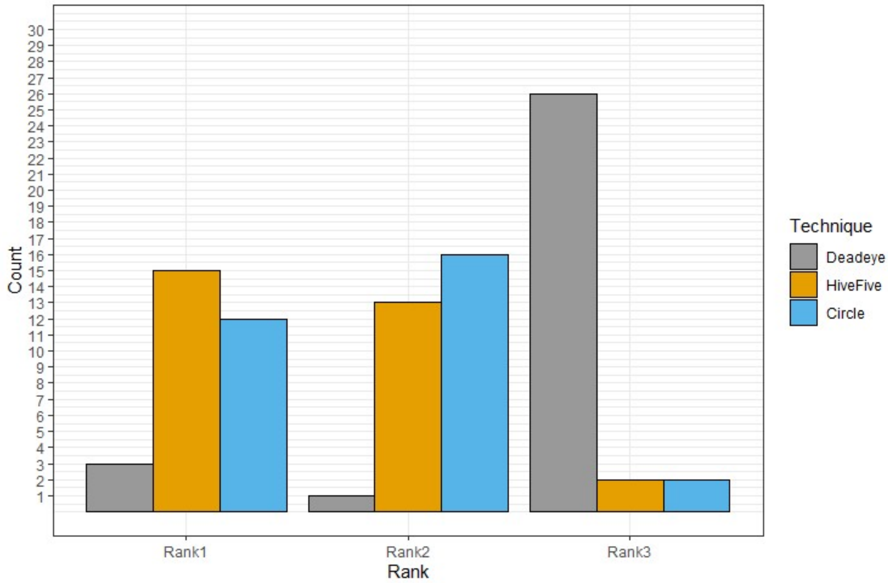


Figure 5.20: Personal ranking of techniques by the participants.

6 Conclusion and Outlook

Many researchers already studied visual attention and how to direct it within visualisations and graphical representations [HE12; MK09; WH04]. While most of the previous work considered the attention guidance within 2D representations [BMSG09; WH04], there is only few work evaluating techniques for AR and VR [GSEM17; LSGB20]. Since VR became more popular for gaming and 360° movies, there is an increasing need for efficient and immersive visual attention techniques [RBH19; RH18]. However, the majority of attention guidance research studies such techniques using rather static environments without including any other, potentially disturbing, objects [GTA+19; LSGB20].

For this reason, this work investigated the influence of static and dynamic distractors on common visual attention techniques in VR. Therefore, we compared the techniques Deadeye, HiveFive and Circle within a virtual nature environment using quantitative measures as well as subjective task load information. We included two different types of distractors. For the static distractors, we applied non-moving and similar looking objects whereas the dynamic distractors included also common, moving objects (e.g, bees or birds) that may appear within similar real world environments.

6.1 Design Implications

Although we cannot confirm H1, we saw some effects between certain distractorlevel for all techniques (e.g. for the Circle between the static and dynamic distractorlevel). Hence, we assume that distractor objects have an influence on the perception of visual attention techniques. This effect is also visible when considering the trend for our questions regarding the distraction and difficulty by the scene elements. Based on the qualitative data, we also assume that there are differences between certain distractors (e.g., bees are more eye-catching and distracting than the pollen for HiveFive) since some distractors were named more often in the comments. Therefore, we suggest further work considering the general effect of distractors as well as the type of distractor that might trigger it.

For Deadeye, we can conclude that it generally needs more learning time compared to HiveFive and Circle. It mostly resulted in the highest scores when looking at the quantitative measures (i.e., completion time) but also for the subscales of the subjective task load as we can confirm H3. However, the mean of the completion time decreases when looking at the values for the consecutive runs. This also aligns with the further comments since several participants argued that the technique got easier after a few targets. Hence, we suspect that Deadeye can be used for general highlighting if participants get an instruction on its behaviour. However, due to its subtle properties and our findings that for an increase in distractorlevel it is more difficult to perceive, we assume it is rather unsuitable for attention guidance in CVR contexts when there is a lot going on. Additionally, it is possible that a target might appear behind the viewer in a 360° environment and since Deadeye is even harder to perceive in the peripheral vision, it is very difficult to direct the attention to a position out of vision. Besides this, Deadeye introduces an uncommon and unnatural behaviour to our vision by perceiving a target to one eye and although it does not cause any strains in short-term usage, there is still no data regarding long-term usage.

Moreover, HiveFive performed very well for all distractorlevel even for the dynamic distractorlevel which included distractor objects with similar properties since we cannot confirm H2. However, by considering the results of our own questions and the general comments, we can see that distractors influenced the perception of the swarm. According to the decrease of the values for the third and forth own question, we can see that participants felt distracted by certain distractors. In addition, several participants mentioned that similar looking and behaving distractor objects (e.g., bees) made the task more difficult. Since this aligns with the findings of Olk et al. [ODZK18] that similar properties of targets and distractors influence the perception, we assume this is also the case for visual attention techniques that share properties (colour, motion) with other scene objects. Hence, it might be interesting using a swarm which is more adapted to the overall scene since we only used basic spheres instead of realistic appearing bees. It is also necessary to find a trade-off between the effectiveness of guidance and annoyance within virtual environments as Waldner et al. [WLB+14] already suggested for 2D since several participants stated HiveFive to be stressful and annoying.

Lastly, our rather easy and static technique Circle performed surprisingly well and mostly outperformed both other techniques. It also achieved rather high results for our second own question which referred to the adaptation in the environment. However, we suspect that most of the participants did not interpret fitting well in terms of immersion due to some comments. In general, due to this results it might be reasonable to question why things have to be as authentic as possible in VR. Since VR is built of virtual elements why not using it as an advantage by enhancing this environment with artificial objects to make things easier (e.g., a circle for attention guidance).

6.2 Limitations

In general, our user study findings are limited due to some factors which potentially effect our results. First, since we did not have a normal distribution of our data, it might be reasonable to conduct this user study with more participants and also for a greater age distribution to gather more insights. In addition, as previously mentioned, we randomised our distractorlevel order. However, we used a general randomisation function provided by Unity. Therefore, the randomisation was not balanced in terms of how often a distractor level appeared as first or second or last. Thus, it is possible that a certain distractorlevel occurred more often as first for one technique while being more often the last for another. For this reason, the results might be biased and some effects only appeared or did not appear as a result.

Besides this, several participants stated in the comments that they had struggles with clicking in VR using the motion controllers at first. Although we integrated a training task to avoid such problems, they had to click multiple times until it was recognized as a successful click. While they also mentioned that after a few rounds they could perform better, we assume that this had also an effect on our quantitative measures (i.e., completion time and wrong clicked objects). For this reason, we suggest to use a training task that is more accustomed to the actual task. Even though we did not want to give any hints to potential interactions, it is possible to use a scenario that is similar to the task using other targets (e.g., in our case removing books from a shelf in a similar distance as the target tree).

6.3 Outlook

Since there is only few work considering distractors in VR, more research is needed to gather further insights in terms of potential effects and also its factors. It is also necessary to evaluate certain techniques in further environments or contexts using differing distractors (e.g., within a supermarket, city or museum environment). Hence, it is possible that techniques perform better for some distractors than with others. In this case, it would be also interesting to consider a larger area in which targets could be highlighted since our and most previous work only considered highlighted targets within the field of view. Besides this, it would be also possible to apply attention guidance on actual application fields (e.g., gaming or scientific background) to differentiate if certain techniques outperform in such areas. This is especially useful since not all contexts require the same type of immersion and sense of presence. For this, it might be also interesting to consider long-term effects as we assume it is possible that some techniques perform generally better or worse over a long time. Additionally, while in our user study participants did not have to hurry, there are certain contexts that require a rapid detection of target objects and hence, it is advisable to study certain differences between techniques in terms of time pressure. In this case, distractors could also play an important role as they might make the task more difficult.

In addition, we recognized that some distractors were mentioned quite often while others were never mentioned. Due to this, we assume that some distractors influence the perception more than others. This might be due to the similar properties as it was often mentioned but also due to some other, not found, factors. Hence, it is reasonable to further investigate factors why and how certain distractors distract and make the task more difficult. This might be also interesting in terms of AR. Regarding the usage of distractors, we highlight that there could be a potential relation between the potential masking of distractors and gaming experience. Although we did not measure this, we assume that this provides further insights of the properties of distractors. Besides the gaming experience, we also suspect that age can have an influence on the perception of distractors and targets but also on visual attention guidance in general. However, we did not have enough data and variations in age to be able to make a proper statement. For this reason, we further emphasize that it might be also interesting to consider the influence of long-sighted and short-sighted vision on the perception of subtle gaze guidance techniques. Due to hardware problems it is possible that people that wear glasses might see a blurred virtual environment and hence, subtle techniques might only be barely visible. Even though subtle techniques can be more beneficial for immersion, such implications have to be considered as well.

Besides the rather general future work, there are also additional considerations regarding the evaluation of the techniques. Since more participants stated that they had struggles to perceive Deadeye for a dominant left eye, we assume that there is a difference regarding showing a target on the dominant or non-dominant eye. We also suggest that it might be reasonable to consider our collected data again to quantitatively evaluate the data in terms of eye dominance. In addition, several participants liked HiveFive due to its eye-catching and bright yellow colour. However, HiveFive tries to imitate a swarm and hence, it is advisable to investigate again the perception of it using realistic objects instead of using generic yellow spheres. Besides HiveFive also the Circle had a rather bright and intense colour. In order to preserve the immersion and to enhance a sense of presence, it would be interesting to evaluate how different colour properties influence the perception of visual attention techniques.

In general, it can be seen that visual attention is important in several routines in our daily life but also when looking at data. For the latter, researchers can facilitate and provide assistance to viewers by using efficient visual attention techniques especially for AR and VR. This work shows that more research is required to build efficient and visually appealing attention guidance techniques that work well regardless of the context. While this work provides a first attempt to study the influence of distractors, it is necessary to build upon it to find influencing factors and also to evaluate further techniques. Thus, at some point, it might be possible to drive a virtual car under demanding conditions without feeling stressed.

Bibliography

[amu] amusedART. *Fantasy Bee*. accessed on 10.06.2022. URL: https://assetstore.unity.com/packages/3d/characters/animals/fantasy-bee-135487 (cit. on p. 36).

[BMSG09] R. Bailey, A. McNamara, N. Sudarsanam, C. Grimm. "Subtle Gaze Direction". In: *ACM Trans. Graph.* 28.4 (Sept. 2009). ISSN: 0730-0301. DOI: 10.1145/1559755.1559757. URL: https://doi.org/10.1145/1559755.1559757 (cit. on pp. 21–24, 28, 79).

[Car18] M. Carrasco. "How visual spatial attention alters perception". In: *Cognitive Processing* 19 (July 2018). DOI: 10.1007/s10339-018-0883-4 (cit. on pp. 15, 18).

[CF05] G. Cumming, S. Finch. "Inference by Eye: Confidence Intervals and How to Read Pictures of Data". In: *The American psychologist* 60 (Feb. 2005), pp. 170–80. DOI: 10.1037/0003-066X.60.2.170 (cit. on pp. 52–55, 59, 60).

[CM84] W. S. Cleveland, R. McGill. "Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods". In: *Journal of the American Statistical Association* 79.387 (1984), pp. 531–554. ISSN: 01621459. URL: http://www.jstor.org/stable/2288400 (cit. on pp. 15, 17, 18).

[Cum14] G. Cumming. "The New Statistics: Why and How". In: *Psychological Science* 25.1 (2014). PMID: 24220629, pp. 7–29. DOI: 10.1177/0956797613504966. eprint: https://doi.org/10.1177/0956797613504966. URL: https://doi.org/10.1177/0956797613504966 (cit. on pp. 52–55, 59, 60).

[DBD07] M. Das, D. Bennett, G. Dutton. "Visual attention as an important visual function: an outline of manifestations, diagnosis and management of impaired visual attention". In: *The British journal of ophthalmology* 91 (Dec. 2007), pp. 1556–60. DOI: 10.1136/bjo.2006.104844 (cit. on p. 18).

[Dem] D. Demchenko. *Simple Office*. accessed on 10.06.2022. URL: https://assetstore.unity.com/packages/3d/props/interior/simple-office-28730#description (cit. on pp. 30, 39).

[din] dinopunch. *Living Birds*. accessed on 10.06.2022. URL: https://assetstore.unity.com/packages/3d/characters/animals/birds/living-birds-15649 (cit. on p. 36).

[Fou] T. R. Foundation. *The R Project for Statistical Computing*. accessed on 10.06.2022. URL: https://www.r-project.org/ (cit. on p. 51).

[Gama] GameDevChef. *Flocking AI Algorithm in Unity Tutorial*. accessed on 10.06.2022. URL: https://www.youtube.com/watch?v=mBVarJm3Tgk (cit. on p. 34).

[Gamb] P. Games. *Realistic Tree 9 [Rainbow Tree]*. accessed on 10.06.2022. URL: https://assetstore.unity.com/packages/3d/vegetation/trees/realistic-tree-9-rainbow-tree-54622 (cit. on pp. 31, 36).

[GAM18] S. Grogorick, G. Albuquerque, M. Maqnor. “Gaze Guidance in Immersive Environments”. In: *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 2018, pp. 563–564. DOI: 10.1109/VR.2018.8446215 (cit. on pp. 15, 23, 25).

[Gmb] I. GmbH. *Butterfly (Animated)*. accessed on 10.06.2022. URL: https://assetstore.unity.com/packages/3d/characters/animals/insects/butterfly-animated-58355 (cit. on p. 37).

[Gre20] H. Greutman. *What is Visual Perception?* accessed on 10.06.2022. Mar. 2020. URL: https://www.growinghandsonkids.com/what-is-visual-perception.html (cit. on p. 17).

[GSEM17] S. Grogorick, M. Stengel, E. Eisemann, M. Magnor. “Subtle Gaze Guidance for Immersive Environments”. In: *Proceedings of the ACM Symposium on Applied Perception*. SAP ’17. Cottbus, Germany: Association for Computing Machinery, 2017. ISBN: 9781450351485. DOI: 10.1145/3119881.3119890. URL: https://doi.org/10.1145/3119881.3119890 (cit. on pp. 15, 23, 24, 28, 79).

[GTA+19] S. Grogorick, J.-P. Tauscher, G. Albuquerque, M. Kassubeck, M. Magnor. “Towards VR Attention Guidance: Environment-dependent Perceptual Threshold for Stereo Inverse Brightness Modulation”. In: *Proc. ACM Symposium on Applied Perception (SAP)*. Sept. 2019. DOI: 10.1145/3343036.3343137 (cit. on pp. 15, 23–26, 79).

[HBB20] P. Hein, M. Bernhagen, A. C. Bullinger. “Two is Better Than One. Improved Attention Guiding in AR by Combining Techniques”. In: *IEEE Computer Graphics and Applications* 40.5 (2020), pp. 57–66. DOI: 10.1109/MCG.2020.3012274 (cit. on p. 26).

[HE12] C. G. Healey, J. T. Enns. “Attention and Visual Memory in Visualization and Computer Graphics”. In: *IEEE Transactions on Visualization and Computer Graphics* 18 (2012), pp. 1170–1188 (cit. on pp. 15, 17, 19–22, 28, 31, 79).

[HLB+16] H. Huang, N.-C. Lin, L. Barrett, D. Springer, H.-C. Wang, M. Pomplun, L.-F. Yu. “Analyzing visual attention via virtual environments”. In: Nov. 2016, pp. 1–2. DOI: 10.1145/2992138.2992152 (cit. on p. 23).

[Inc22] C. Inc. *Visual Perception Cognitive Ability-Neuropsychology*. accessed on 10.06.2022. 2022. URL: https://www.cognifit.com/science/cognitive-skills/visual-perception (cit. on pp. 15, 17).

[KCWK19] A. Krekhov, S. Cmentowski, A. Waschk, J. Krüger. “Deadeye Visualization Revisited: Investigation of Preattentiveness and Applicability in Virtual Environments”. In: *IEEE transactions on visualization and computer graphics* PP (Aug. 2019). DOI: 10.1109/TVCG.2019.2934370 (cit. on p. 32).

[KK19] A. Krekhov, J. H. Krüger. “Deadeye: A Novel Preattentive Visualization Technique Based on Dichoptic Presentation”. In: *IEEE Transactions on Visualization and Computer Graphics* 25 (2019), pp. 936–945 (cit. on pp. 15, 19, 22, 25, 31, 32, 45, 74).

[Lag] S. Lague. *Bézier Path Creator*. accessed on 10.06.2022. URL: https://assetstore.unity.com/packages/tools/utilities/b-zier-path-creator-136082 (cit. on p. 37).

[LF98] S. J. Luck, M. A. Ford. “On the role of selective attention in visual perception”. In: *Proceedings of the National Academy of Sciences* 95.3 (1998), pp. 825–830. DOI: 10.1073/pnas.95.3.825. eprint: https://www.pnas.org/doi/pdf/10.1073/pnas.95.3.825. URL: https://www.pnas.org/doi/abs/10.1073/pnas.95.3.825 (cit. on pp. 15, 18).

[LSGB20] D. Lange, T. C. Stratmann, U. Gruenefeld, S. Boll. “HiveFive: Immersion Preserving Attention Guidance in Virtual Reality”. In: New York, NY, USA: Association for Computing Machinery, 2020, pp. 1–13. ISBN: 9781450367080. URL: https://doi.org/10.1145/3313831.3376803 (cit. on pp. 15, 23, 24, 26, 28, 30, 33, 34, 37, 40, 43, 45, 73, 79).

[Max22] Max-Planck-Gesellschaft. *Visual Perception Psychophysics, Physiology and fMRI Studies*. accessed on 10.06.2022. 2022. URL: https://www.kyb.tuebingen.mpg.de/63265/visual-perception (cit. on pp. 15, 17).

[MBG09] A. McNamara, R. Bailey, C. Grimm. “Search Task Performance Using Subtle Gaze Direction with the Presence of Distractions”. In: *ACM Trans. Appl. Percept.* 6.3 (Sept. 2009). ISSN: 1544-3558. DOI: 10.1145/1577755.1577760. URL: https://doi.org/10.1145/1577755.1577760 (cit. on pp. 28, 36).

[McL18] S. A. McLeod. *Visual perception theory*. Simply Psychology. accessed on 10.06.2022. 2018. URL: www.simplypsychology.org/perception-theories.html (cit. on pp. 15, 17).

[MF95] R. Minghim, A. R. Forrest. “An Illustrated Analysis of Sonification for Scientific Visualisation”. In: *Proceedings of the 6th Conference on Visualization ’95*. VIS ’95. USA: IEEE Computer Society, 1995, p. 110. ISBN: 0818671874 (cit. on p. 23).

[Mica] Microsoft. *Microsoft Documentation*. accessed on 10.06.2022. URL: https://docs.microsoft.com/de-de/dotnet/api/system.collections.generic.dictionary-2?view=net-6.0 (cit. on p. 35).

[Micb] Microsoft. *Mixed Reality Toolkit*. accessed on 10.06.2022. URL: https://docs.microsoft.com/de-de/windows/mixed-reality/mrtk-unity/?view=mrtkunity-2021-05 (cit. on p. 29).

[MK09] S. A. McMains, S. Kastner. “Visual Attention”. In: *Encyclopedia of Neuroscience*. Ed. by M. D. Binder, N. Hirokawa, U. Windhorst. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 4296–4302. ISBN: 978-3-540-29678-2. DOI: 10.1007/978-3-540-29678-2_6344. URL: https://doi.org/10.1007/978-3-540-29678-2_6344 (cit. on pp. 15, 18, 19, 79).

[NMH+16] L. T. Nielsen, M. B. Møller, S. D. Hartmeyer, T. C. M. Ljung, N. C. Nilsson, R. Nordahl, S. Serafin. “Missing the Point: An Exploration of How to Guide Users’ Attention during Cinematic Virtual Reality”. In: *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*. VRST ’16. Munich, Germany: Association for Computing Machinery, 2016, pp. 229–232. ISBN: 9781450344913. DOI: 10.1145/2993369.2993405. URL: https://doi.org/10.1145/2993369.2993405 (cit. on pp. 23, 26, 43).

[ODZK18] B. Olk, A. Dinu, D. Zielinski, R. Kopper. “Measuring visual search and distraction in immersive virtual reality”. English. In: *ROY SOC OPEN SCI* 2018.5 (Mar. 2018). ISSN: 2054-5703 (cit. on pp. 27, 36, 37, 80).

[POT08] Y. Pinto, C. Olivers, J. Theeuwes. "Selecting from dynamic environments: Attention distinguishes between blinking and moving". In: *Perception and psychophysics* 70 (Feb. 2008), pp. 166–78. DOI: 10.3758/PP.70.1.166 (cit. on pp. 22, 28).

[RBH19] S. Rothe, D. Buschek, H. Hussmann. "Guidance in Cinematic Virtual Reality-Taxonomy, Research Status and Challenges". In: *Multimodal Technologies and Interaction* 3 (Mar. 2019), p. 19. DOI: 10.3390/mti3010019 (cit. on pp. 15, 26, 43, 79).

[Rey87] C. W. Reynolds. "Flocks, Herds and Schools: A Distributed Behavioral Model". In: *SIGGRAPH Comput. Graph.* 21.4 (Aug. 1987), pp. 25–34. ISSN: 0097-8930. DOI: 10.1145/37402.37406. URL: https://doi.org/10.1145/37402.37406 (cit. on pp. 33, 34, 37, 67).

[RH18] S. Rothe, H. Hußmann. "Guiding the Viewer in Cinematic Virtual Reality by Diegetic Cues". In: *Augmented Reality, Virtual Reality, and Computer Graphics*. Ed. by L. T. De Paolis, P. Bourdot. Cham: Springer International Publishing, 2018, pp. 101–117. ISBN: 978-3-319-95270-3 (cit. on pp. 15, 23, 26, 79).

[RP17] P. Renner, T. Pfeiffer. "Attention guiding techniques using peripheral vision and eye tracking for feedback in augmented-reality-based assistance systems". In: *2017 IEEE Symposium on 3D User Interfaces (3DUI)*. 2017, pp. 186–194. DOI: 10.1109/3DUI.2017.7893338 (cit. on p. 26).

[RP18] P. Renner, T. Pfeiffer. "Attention Guiding Using Augmented Reality in Complex Environments". In: *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 2018, pp. 771–772. DOI: 10.1109/VR.2018.8446396 (cit. on p. 26).

[RP20] P. Renner, T. Pfeiffer. "AR-Glasses-Based Attention Guiding for Complex Environments: Requirements, Classification and Evaluation". In: *Proceedings of the 13th ACM International Conference on PErvasive Technologies Related to Assistive Environments*. PETRA '20. Corfu, Greece: Association for Computing Machinery, 2020. ISBN: 9781450377737. DOI: 10.1145/3389189.3389198. URL: https://doi.org/10.1145/3389189.3389198 (cit. on p. 26).

[RSt] RStudio. *RStudio Homepage*. accessed on 10.06.2022. URL: https://www.rstudio.com/ (cit. on p. 51).

[SFR01] T. Schubert, F. Friedmann, H. Regenbrecht. "The Experience of Presence: Factor Analytic Insights". In: *Presence* 10 (June 2001), pp. 266–281. DOI: 10.1162/105474601300343603 (cit. on p. 23).

[Stu] G. A. Studio. *Ornamental Flower Set*. accessed on 10.06.2022. URL: https://assetstore.unity.com/packages/3d/vegetation/flowers/ornamental-flower-set-11920 (cit. on p. 36).

[Teca] U. Technologies. *Unity*. accessed on 10.06.2022. URL: https://unity.com/ (cit. on pp. 29, 30, 32).

[Tecb] U. Technologies. *Windridge City*. accessed on 10.06.2022. URL: https://assetstore.unity.com/packages/3d/environments/roadways/windridge-city-132222 (cit. on p. 30).

[Wal] D. Wallace. *Simple Collectibles Pack*. accessed on 10.06.2022. URL: https://assetstore.unity.com/packages/3d/props/simple-collectibles-pack-123092 (cit. on p. 35).

[WFGH11] J. O. Wobbrock, L. Findlater, D. Gergle, J. J. Higgins. "The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '11. Vancouver, BC, Canada: Association for Computing Machinery, 2011, pp. 143–146. ISBN: 9781450302289. DOI: 10.1145/1978942.1978963. URL: https://doi.org/10.1145/1978942.1978963 (cit. on pp. 52, 53, 59).

[WH04] J. Wolfe, T. Horowitz. "What Attributes Guide the Deployment of Visual Attention and How Do They Do It?" In: *Nature reviews. Neuroscience* 5 (July 2004), pp. 495–501. DOI: 10.1038/nrn1411 (cit. on pp. 15, 22, 28, 34, 79).

[Wil45] F. Wilcoxon. "Individual Comparisons by Ranking Methods". In: *Biometrics Bulletin* 1.6 (1945), pp. 80–83. ISSN: 00994987. URL: http://www.jstor.org/stable/3001968 (visited on 06/14/2022) (cit. on pp. 53–55, 60, 61).

[WLB+14] M. Waldner, M. Le Muzic, M. Bernhard, W. Purgathofer, I. Viola. "Attractive Flicker — Guiding Attention in Dynamic Narrative Visualizations". In: *IEEE Transactions on Visualization and Computer Graphics* 20.12 (2014), pp. 2456–2465. DOI: 10.1109/TVCG.2014.2346352 (cit. on pp. 15, 22, 28, 80).

[WSA20] M. Whitlock, S. Smart, D. Albers Szafir. "Graphical Perception for Immersive Analytics". In: Mar. 2020, pp. 616–625. DOI: 10.1109/VR46266.2020.00084 (cit. on p. 18).

[WU18] J. Wolfe, I. Utochkin. "What is a preattentive feature?" In: *Current Opinion in Psychology* 29 (Nov. 2018). DOI: 10.1016/j.copsyc.2018.11.005 (cit. on pp. 22, 34).

[Yu19] K. A. Yu. "The Library: A Non-Intrusive Gaze Directed Virtual Reality Animation". In: *2019 IEEE 2nd Workshop on Animation in Virtual and Augmented Environments (ANIVAE)*. 2019, pp. 1–4. DOI: 10.1109/ANIVAE47543.2019.9050930 (cit. on p. 26).

All links were last followed on June 10, 2022.

A Visual Disability

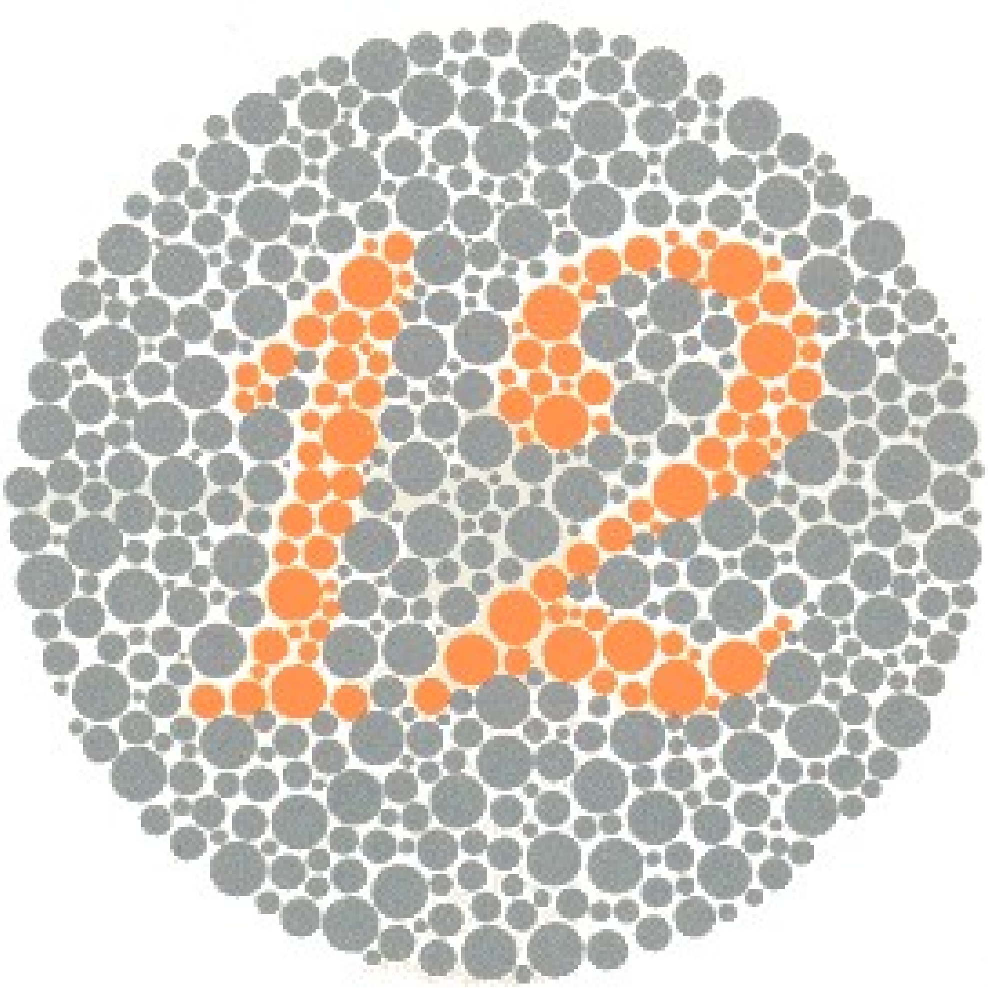


Figure A.1: Ishihara Table 12.

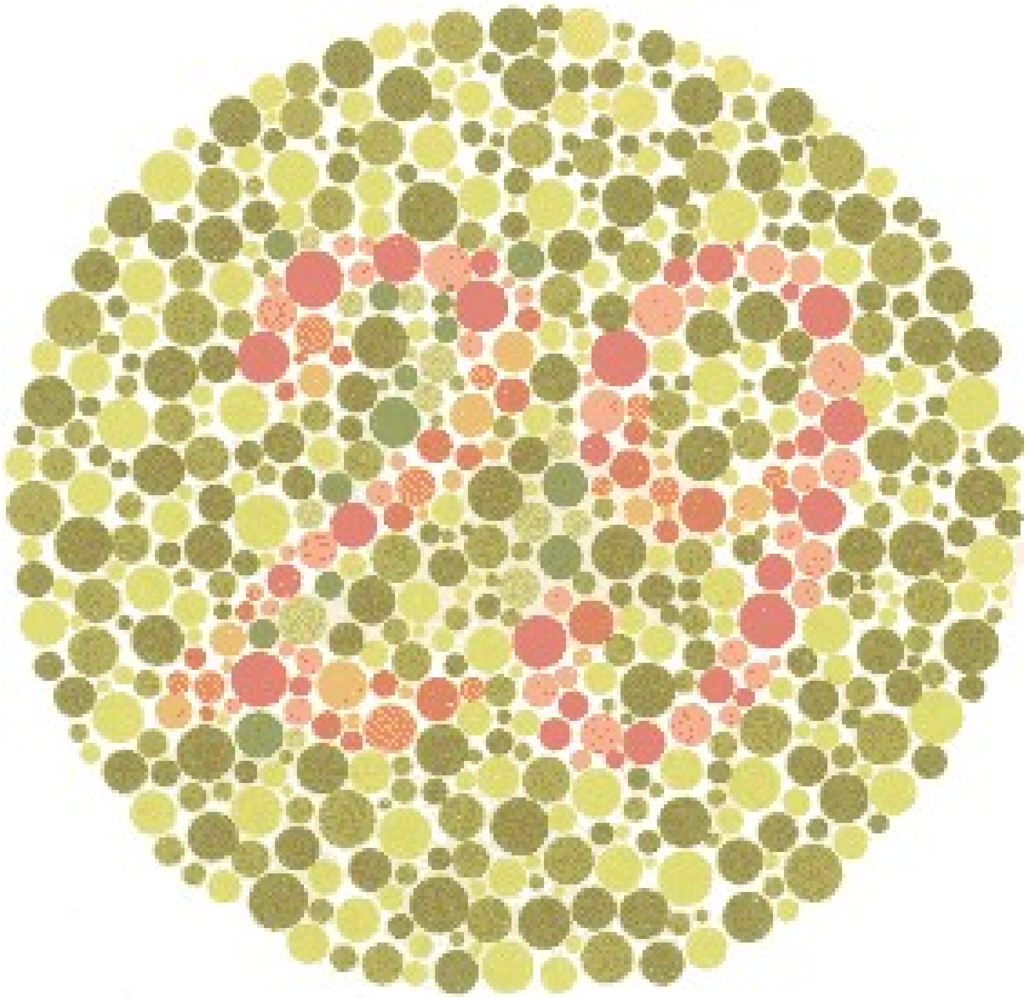


Figure A.2: Ishihara Table 29.

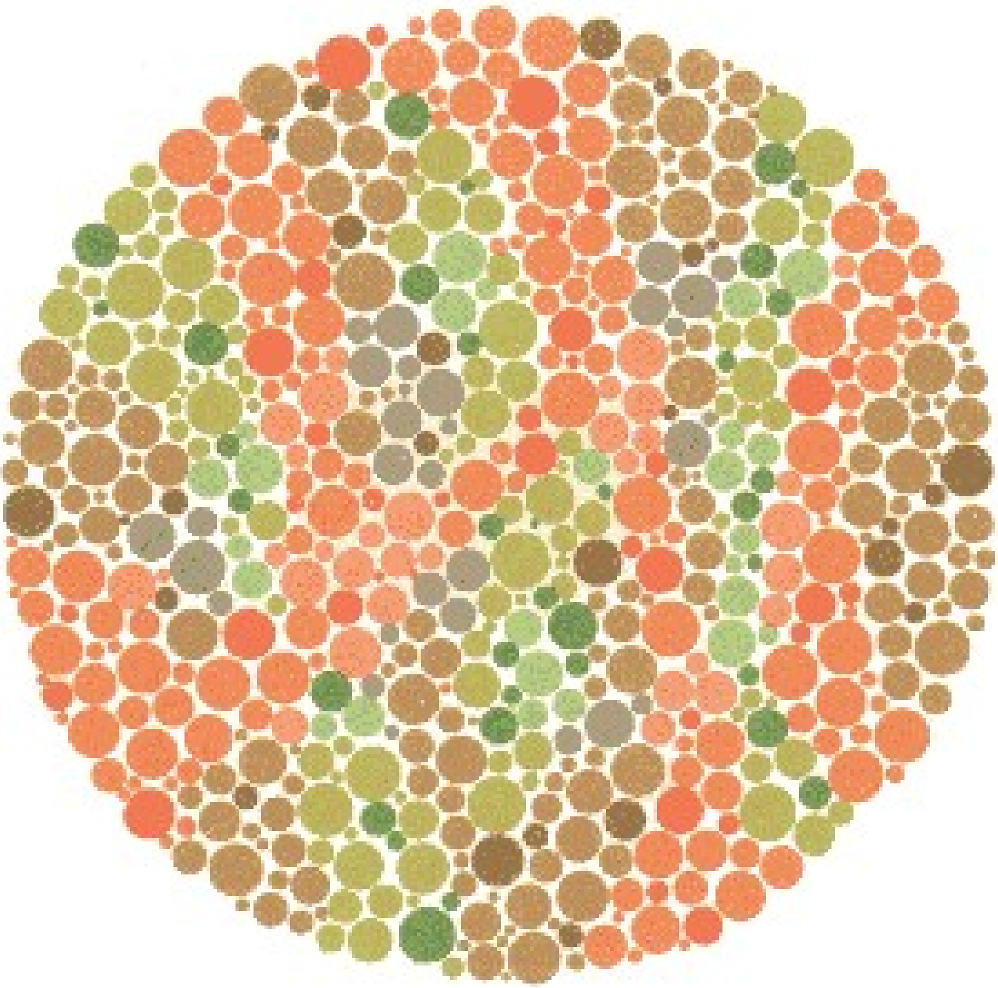


Figure A.3: Ishihara Table No.

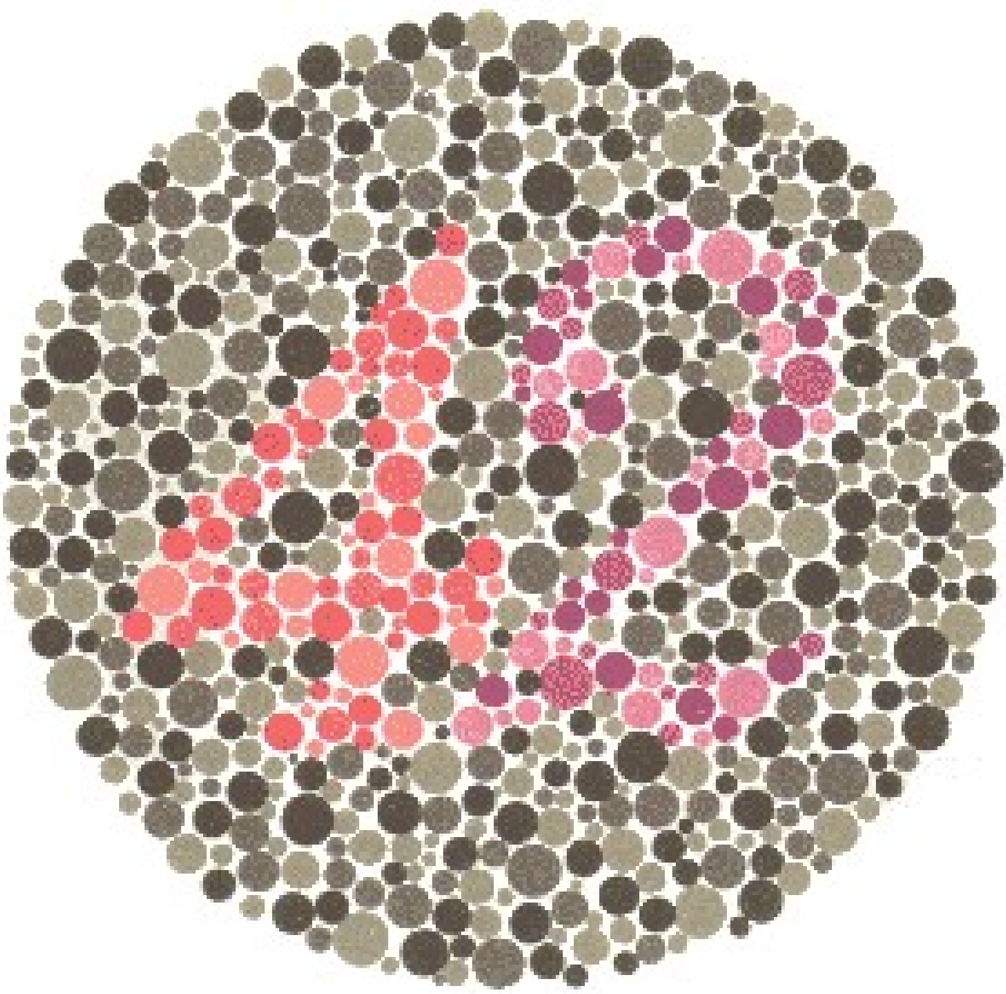


Figure A.4: Ishihara Table 42.

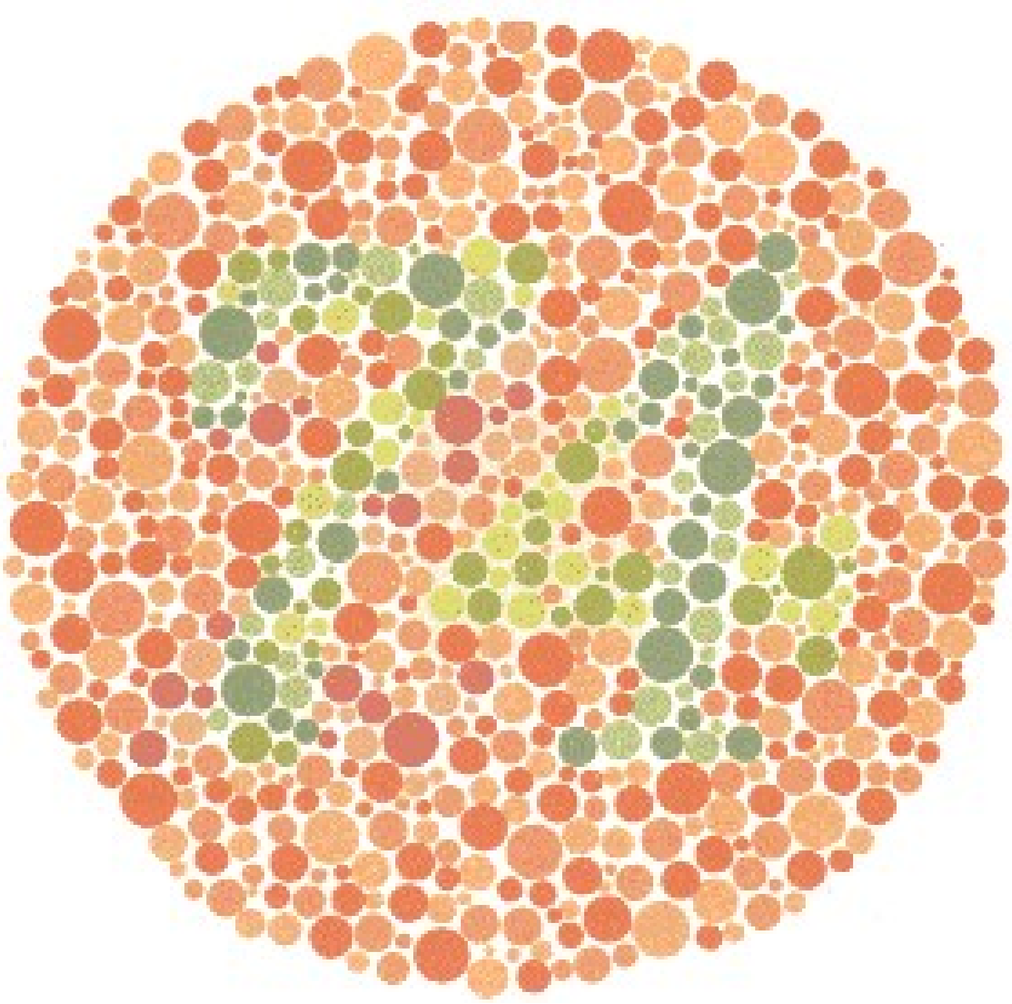


Figure A.5: Ishihara Table 74.

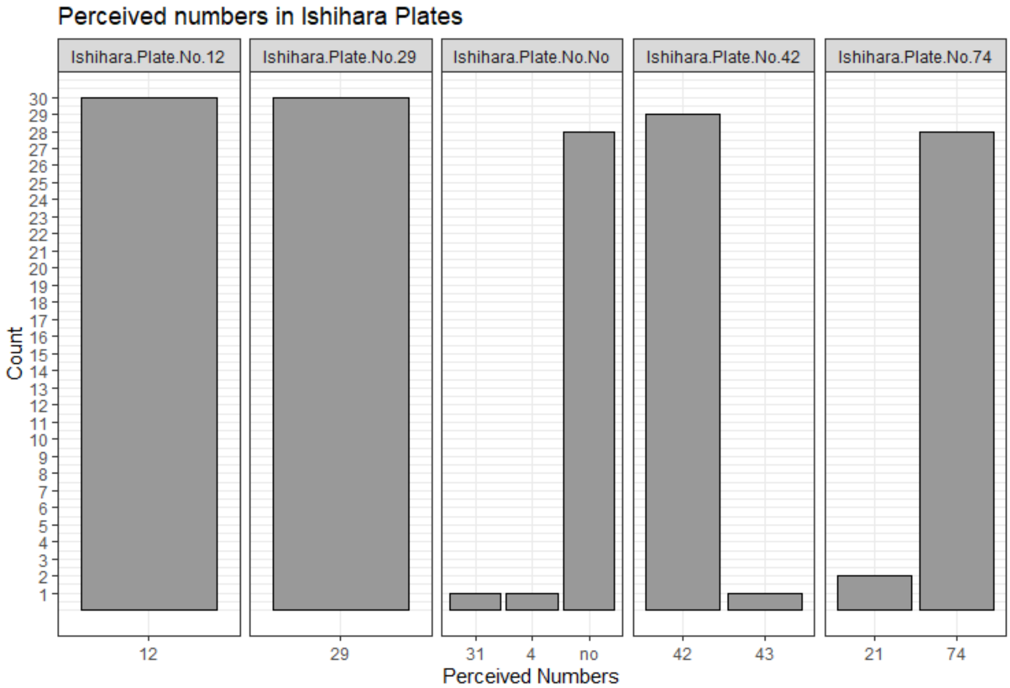


Figure A.6: Ishihara Plot.

B Questionnaire of the User Study

LimeSurvey

Thank you! Your participation matters for a great outcome of my Thesis. You can now start with the first section :)

Section A: Demographic Questionnaire

A1. Please enter your Participant ID

Your participant ID is written on your Consent Form and Privacy Form. You can also ask the study coordinator.

A2. Please provide your age

- 18-21
- 22-25
- 26-30
- 31-35
- 36-40
- 41-45
- 46-50
- 51-55
- 56-60
- 61 and over

A3. Please provide your gender.

- Male
- Female
- Other
- No answer
- Prefer to self-describe

Prefer to self-describe

LimeSurvey

A4. **Are you short-sighted?**

Yes ☐

No ☐

A5. **Are you long-sighted?**

Yes ☐

No ☐

A6. **Do you have any other visual disabilities (e.g., Red-Green Weakness, Grauer Star)**

If yes, please provide in the commentar section.

Yes ☐

No ☐

A7. **Which eye is your dominant eye?**

If you don't know, please ask the study coordinator for further instructions.

Left Eye ☐

Right Eye ☐

A8. **Please provide your personal rating under the following categories.**

	0 - Never	1	2	3	4 - Everyday
Virtual Reality Experience	☐	☐	☐	☐	☐
(PC-) Gaming Usage	☐	☐	☐	☐	☐

Section B: Visual Disability Questionnaire

B1. **What number do you see? (If you don't see anything, write "no")**

LimeSurvey

B2. **What number do you see? (If you don't see anything, write "no")**

B3. **What number do you see? (If you don't see anything, write "no")**

B4. **What number do you see? (If you don't see anything, write "no")**

B5. **What number do you see? (If you don't see anything, write "no")**

Section C: Condition Selection (1-1)

This part will be filled out by the study coordinator.

C1. **Please provide the first technique**

- Deadeye ☐
- HiveFive ☐
- Circle ☐

C2. **Please provide the first Condition.**

- No ☐
- Static ☐
- Dynamic ☐

LimeSurvey

Section D: First Technique - Part 1

D1. Please provide your opinion for the following questions.

The numbers below have the following meaning:

Mental / Physical / Temporal Demand, Effort, Frustration (0 = Very Low | 10 = Neutral | 20 = Very High) Performance (0 = Failure | 10 = Neutral | 20 = Perfect)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Me D nd ow nta demanding was the task?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
ysic Der d (v p ica demanding was the task?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Te ora em l (H hu ed ash was the pace of the task?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
for nce ow ce l w yo ac npl ng at you were asked to do?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Ef (H ha did h to k t co ish ur level of performance?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Fr atio Ho nse e, c our d, ate tre l and annoyed were you?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

D2. Please provide your opinion on the following statements

	strongly disagree	disagree	neutral	agree	strongly agree
I could perceive the highlighted object well	☐	☐	☐	☐	☐
The technique fits well into the overall scene setting.	☐	☐	☐	☐	☐
The current visual scene elements distracted me from the target search	☐	☐	☐	☐	☐
The current visual scene elements made the task more difficult	☐	☐	☐	☐	☐

D3. Is there anything else you want to mention?

LimeSurvey

Section E: Condition Selection (1-2)

This part will be filled out by the study coordinator.

E1. Please provide the second Condition.

- No ☐
- Static ☐
- Dynamic ☐

Section F: First Technique - Part 2

F1. Please provide your opinion for the following questions.

The numbers below have the following meaning:

Mental / Physical / Temporal Demand, Effort, Frustration (0 = Very Low | 10 = Neutral | 20 = Very High) Performance (0 = Failure | 10 = Neutral | 20 = Perfect)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Mental Demand (How mentally demanding was the task?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Physical Demand (How physically demanding was the task?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Temporal Demand (How hurried or rushed was the pace of the task?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Performance (How successful were you in accomplishing what you were asked to do?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Effort (How hard did you have to work to accomplish your level of performance?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Frustration (How insecure, discouraged, irritated, stressed and annoyed were you?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

F2. Please provide your opinion on the following statements

	strongly disagree	disagree	neutral	agree	strongly agree
I could perceive the highlighted object well	☐	☐	☐	☐	☐
The technique fits well into the overall scene setting.	☐	☐	☐	☐	☐
The current visual scene elements distracted me from the target search	☐	☐	☐	☐	☐
The current visual scene elements made the task more difficult	☐	☐	☐	☐	☐

LimeSurvey

F3. **Is there anything else you want to mention?**

Section G: Condition Selection (1-3)

This part will be filled out by the study coordinator.

G1. **Please provide the third Condition.**

- No ☐
- Static ☐
- Dynamic ☐

Section H: First Technique - Part 3

H1. **Please provide your opinion for the following questions.**

The numbers below have the following meaning:

Mental / Physical / Temporal Demand, Effort, Frustration (0 = Very Low | 10 = Neutral | 20 = Very High) Performance (0 = Failure | 10 = Neutral | 20 = Perfect)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Me D nd w nta demanding was the task?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
ysic Der d (v p ica demanding was the task?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Te ora em l (H hu ed ash vas the pace of the task?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
for nce ow ce l w yo ac npl ng at you were asked to do?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Ef (H ha did h to k t co ish ir level of performance?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Fr ati Ho nse e, ou d, ate tre l and annoyed were you?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

LimeSurvey

H2. Please provide your opinion on the following statements

	strongly disagree	disagree	neutral	agree	strongly agree
I could perceive the highlighted object well	☐	☐	☐	☐	☐
The technique fits well into the overall scene setting.	☐	☐	☐	☐	☐
The current visual scene elements distracted me from the target search	☐	☐	☐	☐	☐
The current visual scene elements made the task more difficult	☐	☐	☐	☐	☐

H3. Is there anything else you want to mention?

Section I: Condition Selection (2-1)

This part will be filled out by the study coordinator.

I1. Please provide the second technique

- Deadeye ☐
- HiveFive ☐
- Circle ☐

I2. Please provide the first Condition.

- No ☐
- Static ☐
- Dynamic ☐

LimeSurvey

Section J: Second Technique - Part 1

J1. Please provide your opinion for the following questions.

The numbers below have the following meaning:

Mental / Physical / Temporal Demand, Effort, Frustration (0 = Very Low | 10 = Neutral | 20 = Very High) Performance (0 = Failure | 10 = Neutral | 20 = Perfect)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Mental Demand (How mentally demanding was the task?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Physical Demand (How physically demanding was the task?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Temporal Demand (How hurried or rushed was the pace of the task?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Performance (How successful were you in accomplishing what you were asked to do?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Effort (How hard did you have to work to accomplish your level of performance?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Frustration (How insecure, discouraged, irritated, stressed and annoyed were you?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

J2. Please provide your opinion on the following statements

	strongly disagree	disagree	neutral	agree	strongly agree
I could perceive the highlighted object well	☐	☐	☐	☐	☐
The technique fits well into the overall scene setting.	☐	☐	☐	☐	☐
The current visual scene elements distracted me from the target search	☐	☐	☐	☐	☐
The current visual scene elements made the task more difficult	☐	☐	☐	☐	☐

J3. Is there anything else you want to mention?

LimeSurvey

Section K: Condition Selection (2-2)

This part will be filled out by the study coordinator.

K1. Please provide the second Condition.

- No ☐
- Static ☐
- Dynamic ☐

Section L: Second Technique - Part 2

L1. Please provide your opinion for the following questions.

The numbers below have the following meaning:

Mental / Physical / Temporal Demand, Effort, Frustration (0 = Very Low | 10 = Neutral | 20 = Very High) Performance (0 = Failure | 10 = Neutral | 20 = Perfect)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Mental Demand (How mentally demanding was the task?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Physical Demand (How physically demanding was the task?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Temporal Demand (How hurried or rushed was the pace of the task?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Performance (How successful were you in accomplishing what you were asked to do?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Effort (How hard did you have to work to accomplish your level of performance?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Frustration (How insecure, discouraged, irritated, stressed and annoyed were you?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

L2. Please provide your opinion on the following statements

	strongly disagree	disagree	neutral	agree	strongly agree
I could perceive the highlighted object well	☐	☐	☐	☐	☐
The technique fits well into the overall scene setting.	☐	☐	☐	☐	☐
The current visual scene elements distracted me from the target search	☐	☐	☐	☐	☐
The current visual scene elements made the task more difficult	☐	☐	☐	☐	☐

LimeSurvey

L3. Is there anything else you want to mention?

Section M: Condition Selection (2-3)

This part will be filled out by the study coordinator.

M1. Please provide the third Condition.

- No ☐
- Static ☐
- Dynamic ☐

Section N: Second Technique - Part 3

N1. Please provide your opinion for the following questions.

The numbers below have the following meaning:

Mental / Physical / Temporal Demand, Effort, Frustration (0 = Very Low | 10 = Neutral | 20 = Very High) Performance (0 = Failure | 10 = Neutral | 20 = Perfect)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Mental Demand (How mentally demanding was the task?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Physical Demand (How physically demanding was the task?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Temporal Demand (How hurried or rushed was the pace of the task?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Performance (How successful were you in accomplishing what you were asked to do?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Effort (How hard did you have to work to accomplish your level of performance?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Frustration (How insecure, discouraged, irritated, stressed and annoyed were you?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

LimeSurvey

N2. Please provide your opinion on the following statements

	strongly disagree	disagree	neutral	agree	strongly agree
I could perceive the highlighted object well	☐	☐	☐	☐	☐
The technique fits well into the overall scene setting.	☐	☐	☐	☐	☐
The current visual scene elements distracted me from the target search	☐	☐	☐	☐	☐
The current visual scene elements made the task more difficult	☐	☐	☐	☐	☐

N3. Is there anything else you want to mention?

Section O: Condition Selection (3-1)

This part will be filled out by the study coordinator.

O1. Please provide the third technique

- Deadeye ☐
- HiveFive ☐
- Circle ☐

O2. Please provide the first Condition.

- No ☐
- Static ☐
- Dynamic ☐

Section P: Third Technique - Part 1

P1. Please provide your opinion for the following questions.

The numbers below have the following meaning:

Mental / Physical / Temporal Demand, Effort, Frustration (0 = Very Low | 10 = Neutral | 20 = Very High) Performance (0 = Failure | 10 = Neutral | 20 = Perfect)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Mental Demand (How mentally demanding was the task?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Physical Demand (How physically demanding was the task?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Temporal Demand (How hurried or rushed was the pace of the task?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Performance (How successful were you in accomplishing what you were asked to do?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Effort (How hard did you have to work to accomplish your level of performance?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Frustration (How insecure, discouraged, irritated, stressed and annoyed were you?)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

P2. Please provide your opinion on the following statements

	strongly disagree	disagree	neutral	agree	strongly agree
I could perceive the highlighted object well	☐	☐	☐	☐	☐
The technique fits well into the overall scene setting.	☐	☐	☐	☐	☐
The current visual scene elements distracted me from the target search	☐	☐	☐	☐	☐
The current visual scene elements made the task more difficult	☐	☐	☐	☐	☐

P3. Is there anything else you want to mention?

LimeSurvey

Section Q: Condition Selection (3-2)

This part will be filled out by the study coordinator.

Q1. Please provide the second Condition.

- No ☐
- Static ☐
- Dynamic ☐

Section R: Third Technique - Part 2

R1. Please provide your opinion for the following questions.

The numbers below have the following meaning:

Mental / Physical / Temporal Demand, Effort, Frustration (0 = Very Low | 10 = Neutral | 20 = Very High) Performance (0 = Failure | 10 = Neutral | 20 = Perfect)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Me… D…nd …ow… nta… demanding was the task?)
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	…ysic… Der… d (… v p… ica… demanding was the task?)
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Te… ora… em… l (H… hu… d c… ash… was the pace of the task?)
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	…for… nce… ow… ce… l w… yo… ac… npl… ng… at you were asked to do?)
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Ef… (H… ha… did… h… to… k t… co… ish… ur level of performance?)
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Fr… ati… Ho… nse… e, c… our… d, … ate… tre… l and annoyed were you?)

R2. Please provide your opinion on the following statements

	strongly disagree	disagree	neutral	agree	strongly agree
I could perceive the highlighted object well	☐	☐	☐	☐	☐
The technique fits well into the overall scene setting.	☐	☐	☐	☐	☐
The current visual scene elements distracted me from the target search	☐	☐	☐	☐	☐
The current visual scene elements made the task more difficult	☐	☐	☐	☐	☐

R3. Is there anything else you want to mention?

Section S: Condition Selection (3-3)

This part will be filled out by the study coordinator.

S1. Please provide the third Condition.

- No
- Static
- Dynamic

Section T: Third Technique - Part 3

T1. Please provide your opinion for the following questions.

The numbers below have the following meaning:

Mental / Physical / Temporal Demand, Effort, Frustration (0 = Very Low | 10 = Neutral | 20 = Very High) Performance (0 = Failure | 10 = Neutral | 20 = Perfect)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
																				Me D nd ow nta demanding was the task?)
																				ysic Der d (v p ica demanding was the task?)
																				Te ora em l (H hu ed ash was the pace of the task?)
																				for nce ow ce l w yo ad npl ng at you were asked to do?)
																				Ef (H ha did ha to k t co ish ur level of performance?)
																				Fr ati Ho nse e, ou d, ate tre l and annoyed were you?)

LimeSurvey

T2. Please provide your opinion on the following statements

	strongly disagree	disagree	neutral	agree	strongly agree
I could perceive the highlighted object well	☐	☐	☐	☐	☐
The technique fits well into the overall scene setting.	☐	☐	☐	☐	☐
The current visual scene elements distracted me from the target search	☐	☐	☐	☐	☐
The current visual scene elements made the task more difficult	☐	☐	☐	☐	☐

T3. Is there anything else you want to mention?

Section U: Final Questionnaire

U1. Please provide a rating for the following statements..

	0 - no, not really	1	2	3	4 - yes, strongly
I feel any physical strains now (e.g., headache, eye pain)	☐	☐	☐	☐	☐
I feel any mental strains now (e.g., dizzy)	☐	☐	☐	☐	☐

U2. Which technique did you like most?

Please rank the following techniques based on your own opinion.

Deadeye ☐

HiveFive ☐

Yellow Circle ☐

U3. Please provide an explanation for the ranking of the three techniques created above.

Deadeye

HiveFive

Yellow Circle

LimeSurvey

U4. **Is there anything else you want to mention?**

Again, Thank you for your participation! You have finished now all conditions :)

Declaration

I hereby declare that the work presented in this thesis is entirely my own and that I did not use any other sources and references than the listed ones. I have marked all direct or indirect statements from other sources contained therein as quotations. Neither this work nor significant parts of it were part of another examination procedure. I have not published this work in whole or in part before. The electronic copy is consistent with all submitted copies.

Stuttgart, 15.06.2022 N.Dem

place, date, signature