



Comparison of gaze accuracy and precision in real-world and virtual reality

Stefan Pastel¹ · Chien-Hsi Chen¹ · Luca Martin¹ · Mats Naujoks¹ · Katharina Petri¹ · Kerstin Witte¹

Received: 17 October 2019 / Accepted: 20 May 2020 / Published online: 3 June 2020
© The Author(s) 2020

Abstract

Virtual reality (VR) is popular across many fields and is increasingly used in sports as a training tool. The reason, therefore, is recently improved display technologies, more powerful computation capacity, and lower costs of head-mounted displays for VR. As in the real-world (R), visual effects are the most important stimulus provided by VR. However, it has not been demonstrated whether the gaze behavior would achieve the same level in VR as in R. This information will be important for the development of applications or software in VR. Therefore, several tasks were designed to analyze the gaze accuracy and gaze precision using eye-tracking devices in R and VR. 21 participants conducted three eye-movement tasks in sequence: gaze at static targets, tracking a moving target, and gaze at targets at different distances. To analyze the data, an averaged distance with root mean square was calculated between the coordinates of each target and the recorded gaze points for each task. In gaze accuracy, the results showed no significant differences between R and VR in gaze at static targets (1 m distance, $p > 0.05$) and small significant differences at targets placed at different distances ($p < 0.05$), as well as large differences in tracking the moving target ($p < 0.05$). The precision in VR is significantly worse compared to R in all tasks with static gaze targets ($p < 0.05$). On the whole, this study gives a first insight into comparing foveal vision, especially gaze accuracy and precision between R and VR, and can, therefore, serve as a reference for the development of VR applications in the future.

Keywords Eye-tracking · Virtual reality · Gaze behavior · Head-mounted display · Accuracy · Precision

1 Introduction

Virtual reality (VR) is one of the fast-growing technologies that has many potential applications involving a huge amount of visual cues that are important when analyzing gaze behavior. Currently, the most frequently used VR technology in the field of entertaining or for educational purposes is HMD. The current versions of HMDs have a high-resolution display and are combined with a motion tracking system to ensure high quality of immersion and user experience. Moreover, VR has many advantages, such as the development of highly customizable virtual training

scenes, affordable cost of the system, and high accessibility in a most domestic environments (Düking et al. 2018; Neumann et al. 2018). Moreover, VR can simulate or reproduce images and scenes that are difficult to perform in a real-world scenario. These features make VR an ideal tool for training in different fields, such as rehabilitation (Rose et al. 2005; Duque et al. 2013), health sports (Molina et al. 2014), as well as recreational sports and high-performance sports (Petri et al. 2018a, b).

In addition to the application in healthcare, some studies in the field of sport also showed an improvement after the training sessions using VR, such as in karate (Petri et al. 2019), throwing dart (Tirp et al. 2015), and baseball batting (Gray 2017). All these sports require a continuous focus on the target. For example, a dart player needs to focus on the targets from a fixed distance. In this case, the depth perception and the sharpness of the targets in VR become highly relevant to the results. However, these studies did not provide further information regarding gaze accuracy and precision in the comparison between real-world (R) and VR. Accuracy and precision are considered the most important

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10055-020-00449-3>) contains supplementary material, which is available to authorized users.

✉ Stefan Pastel
stefan.pastel@ovgu.de

¹ Department of Sports Engineering and Movement Science, Institute III: Sports Science, Otto-Von-Guericke-University, Magdeburg, Germany

parameters for data quality of eye movements (Ooms et al. 2015).

The visual system allows the extraction of valuable information from the environment to complete highly skilled actions. In sporting activities, it is essential to perceive teammates, opponents, one's position, or properties concerning one's surroundings (size, target, etc.). Unfortunately, the foveal vision is restricted to 1° – 2° in the field of view (FOV), which leads to constant eye movements to see sharply and extract detailed information (Vater et al. 2017). Gaze accuracy is defined by the degree of visual angle within this FOV (Krokos et al. 2019). It plays an essential role in examining the interindividual differences in attention span and identifying the key points during observation while learning a new movement. Holmqvist et al. (2015) describe gaze accuracy as the averaged deviation between the position of a considered point (target stimulus) and the position captured by the eye-tracking system (point of regard). Precision is defined as the ability to reliably reproduce a measurement given a fixating eye (Nyström et al. 2013). While accuracy defines the distance between true and recorded gaze direction, precision refers to how consistent calculated gaze points are, when true gaze direction is constant (Holmqvist et al. 2012). When we consider using VR as a training tool for sports, the user should ideally have almost the same experiences in perception as he would have in real conditions. Furthermore, due to the progress of technical devices, it is possible to have light eye-tracking systems in VR headsets (Clay et al. 2019). Therefore, it is essential to investigate in depth the specific differences between R and VR regarding gaze behavior.

The goal of the current study is to examine whether gaze accuracy and precision in a simulated virtual scenario are comparable to those of the real environmental setup within different gaze tasks.

2 Related work

With eye-tracking systems, the accuracy of a target stimulus can be determined by specifying angular deviations. Many manufacturers specify a deviation of $<0.5^{\circ}$ in their measurement systems (Feit et al. 2017; Nyström et al. 2013). Higher accuracies are found in the center of FOV because the pupils are the largest detected object in size by the integrated eye-tracking cameras when the target is centered in front of the eye-tracking system (Hornof and Halverson 2002).

On the other hand, the level of accuracy is not the only data quality issue affecting the viability of research results. There are many influencing factors, which can result from either technical or non-technical issues, such as the homogeneity of the testing participants (Blignaut and Wium 2013). Another study showed more factors that might have an impact such as different calibration methods, the individual

characteristics of the human eye, the recording time as well as the gaze direction. Additionally, the operator's experience can also affect gaze data such as accuracy and precision (Nyström et al. 2013). Participant-controlled calibration is predestined for better accuracy and precision. This study has also demonstrated that contact lenses, downward-pointing eyelashes, and smaller pupil sizes harm gaze accuracy. Further studies have figured out that the measurement method and different calculations of gaze accuracy in R also have an influence on gaze accuracy (Feit et al. 2017; Holmqvist et al. 2015; Hooge et al. 2018; Nyström et al. 2013). Moreover, also different environments and different measurement systems can have an affect (Feit et al. 2017). Accuracy is a prerequisite for several technological devices. For example, gaze-based communication technologies, where dwell time selection is a common method for interacting with options on a computer-based surface, require high accuracy, too. Studies analyzing the selected target by using the gaze position for physical interactions (e.g. Pfeuffer et al. 2017) showed how important gaze accuracy is in VR.

When classifying the event detections to define fixations or saccades, another important gaze parameter must be considered: precision. Nyström et al. (2013) tested precision with different systems, resulting in values of 0.01° to 0.05° for tower-mounted systems and 0.03° – 1.03° for remote ones. For instance, a high precision must be given when comparing the number of fixations or the fixation area in R with VR. However, no such comparisons exist up to now.

It needs to be considered, that the representation of the environment in VR takes place via an artificial way. In order to evoke a high presence in the virtual world, the VR system has to manipulate the human perception (Dörner et al. 2013). Possible reasons that have an influence on the performances in VR can result from a distortion of the environment, the perceived depth information, or the level of fidelity. To implement sports training in the VR, it is of great importance to figure out how accurately and precisely the participants perceive short appearing stimuli in fixed and movable conditions. In general, previous investigations aimed to compare different VR applications with each other (Krokos et al. 2019). Clay et al. (2019) have described the technical and practical aspects of eye-tracking in VR and gave an overview of different software and hardware solutions. However, a comparison between VR and R regarding gaze behavior has rarely been made. This study's aim is therefore to examine the differences of gaze accuracy and precision between R and VR under controlled testing conditions.

The main goal of this study is to investigate how gaze accuracy and precision would be affected by different kinds of visual stimuli and scenarios (R vs. VR). Since the resolution in VR is significantly lower than the one of the human eyes and latencies can occur, we assume that differences occur between the gaze accuracy and precision in R and

VR. The three-dimensional world that is shown in VR on a display can lead to differences in distance perception (Loomis and Knapp 2003; Messing and Durgin 2005; Renner et al. 2013). Clay et al. (2019) also mentioned the disparity between vergence and focus, since the distance to the display remains the same and therefore eye strain and fatigue can occur. While we use an older VR application (HTC Vive) in this study, we have to consider all these mentioned limitations, since those components may have an impact on gaze measurements. To avoid possible differences by using devices including different technologies (e.g. manufacturer, measurement method via corneal reflection, use of the same algorithms), the older HMD was chosen to have comparable values between the real and virtual measuring technology (see chapter hardware).

3 Methods

To compare gaze accuracy and precision between R and VR, three tasks were designed: (1) static stimuli appearing at four different positions, (2) a stimulus moving across the screen in the form of an infinity loop, and (3) static stimulus presented at different distances in the center of the screen. We have included those due to the confrontation of different stimuli in daily life. All tasks are performed in R and in VR to have comparable results regarding gaze behavior. The experimental setup, protocols, and data analysis are described in the following subsections (see Fig. 3).

3.1 Participants

Twenty-three young sports students (ten females, eleven males) with an average age of 22.6 ± 3.02 were recruited for this study. However, the data recording of two participants was rejected due to a lack of quality and technical problems during the conduction. All participants took part in three tasks which will be presented later. The participants' previous experiences in VR and in eye-tracking studies were noted. The participants were asked whether they had ever taken part in a VR or eye-tracking study, or if they owned VR applications themselves. Six participants stated that they had already gained VR experience, but none of them owned a VR application. Five participants had already participated in eye-tracking studies. Furthermore, related gaming experiences including the type of games and the frequency of gameplay were also noted. Eleven participants regularly played video games ($M=4.58$ h per week, $SD=2.51$). For vision correction, only participants using contact lenses (8 participants were affected) were allowed, because it was not possible to wear the HMD and glasses simultaneously. All participants received the instructions prior to the study and

gave their written consent. The study was approved by the authors' university's ethics committee.

3.2 Experimental setup

3.2.1 Hardware

The experimental setup is shown in Fig. 1. The participants were seated in front of a table. A chin rest was used to support the participant's chin in a comfortable posture. The participant's head was also fixed during the experiment (Clemotte et al. 2014; Reichert 2019). The height of the center point of the monitor was adjusted to the eye level of each participant (Ooms et al. 2015).

In the real-world testing condition, binocular Eye Tracking Glasses 2.0 (SensoMotoric Instruments, Germany) with a resolution of 1280×960 pixels and the sampling frequency of 60 Hz was used to track the eye movement. A laptop (Lenovo, China) was used to record the eye-tracking data. A 23.5-inch monitor with the resolution of 1920×1080 pixel and 60 Hz refreshing rate (EIZO ColorEdge CG248, Japan) was used to display instructions for the experiment. In this study, the optimal visual acuity distance (between the participant and the monitor) of 1 m was chosen. The distance ensured that the participants could achieve a complete view of the monitor without head movements.

The setup in the VR was the same as in the real-world condition. An HTC Vive HMD (HTC, Taiwan) with an integrated eye-tracking system (Sensorimotor Instrument, SMI, Germany; resolution: 2160×1200 pixels; frequency: 90 Hz; 110° field of view) was used to display the virtual environment. The approximate resolution of the screen that is rendered in VR condition was 720×400 pixels. This VR setup ran on a computer with Intel(R) Core(TM) i7-7700 CPU @ 3.60 GHz, 16 GB RAM, and an NVIDIA GTX 1080 graphics card. The manufacturer specifies a gaze accuracy of 0.4° – 0.5° (SensoMotoric Instruments 2016) overall distances and guarantees parallax compensation (iViewETG User Guide Version 2.7 2016). The precision values were not provided for the mobile system. For the SMI RED 250, which also used the corneal reflection method, a precision of 0.03° is mentioned by the manufacturer (SensoMotoric Instruments 2016).

The following data were recorded by the eye-tracking software: interpupillary distance (IPD), the points of regard of the individual eyes (POR), and the gaze direction vectors of both eyes (Fig. 2).

The stimuli were presented via a PowerPoint presentation, ensuring the same chronological sequence for all participants. Studies have shown that too bright background light can affect gaze accuracy negatively (Drewes et al. 2011). In the current study, we therefore chose a gray screen background. The fixation cross (in the middle

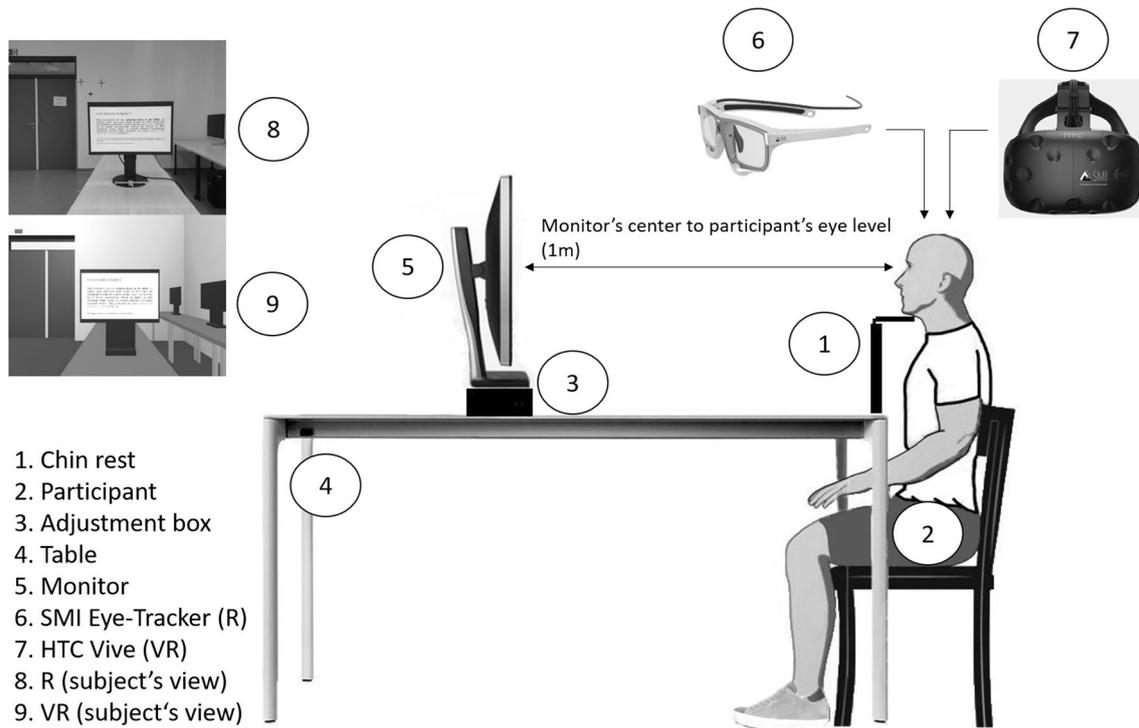


Fig. 1 Overview of the experimental setup

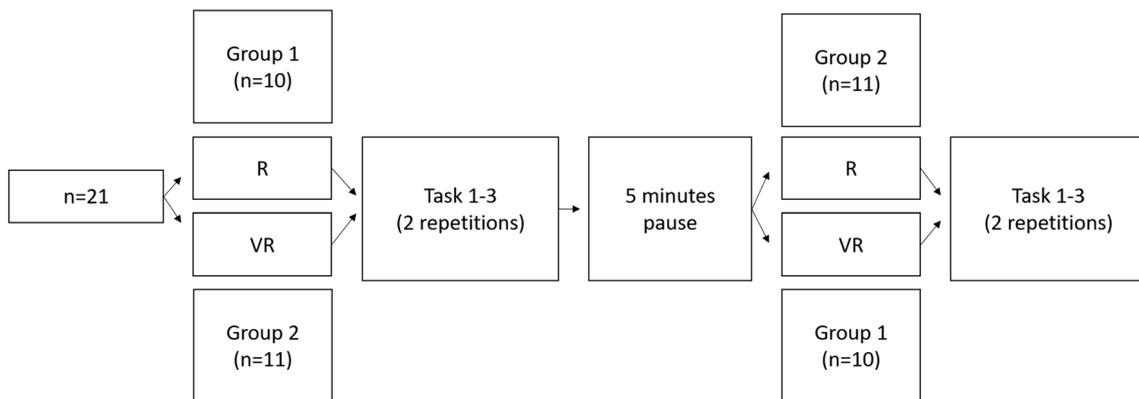


Fig. 2 Overview of the experimental conduction. Each task was performed twice in R and VR. The participants had to complete all tasks in both conditions. For detailed information on tasks 1–3, see Fig. 3

of the screen) and the stimuli at the corners were 15.64° (FOV) apart. The presented crosses for task 1 were 2 cm wide and 2.5 cm high (0.99° horizontal and 1° vertical on the FOV), the diameter of the dot in task 2 was 1 cm (0.57° on the FOV), and the cross in task 3 was 8 cm wide and high (the cross in the middle of the white one was 1 cm wide and high, again 0.57° on the FOV, see Fig. 3 Task

3). The presentation of the visual stimuli was the same in both conditions. For further information see also Fig. 4.

3.2.2 Software

For data recording and extraction in R, the iViewETG 2.7 and BeGaze 3.6 (SensoMotoric Instruments, Germany, 2009) were used. The VR environment was created within

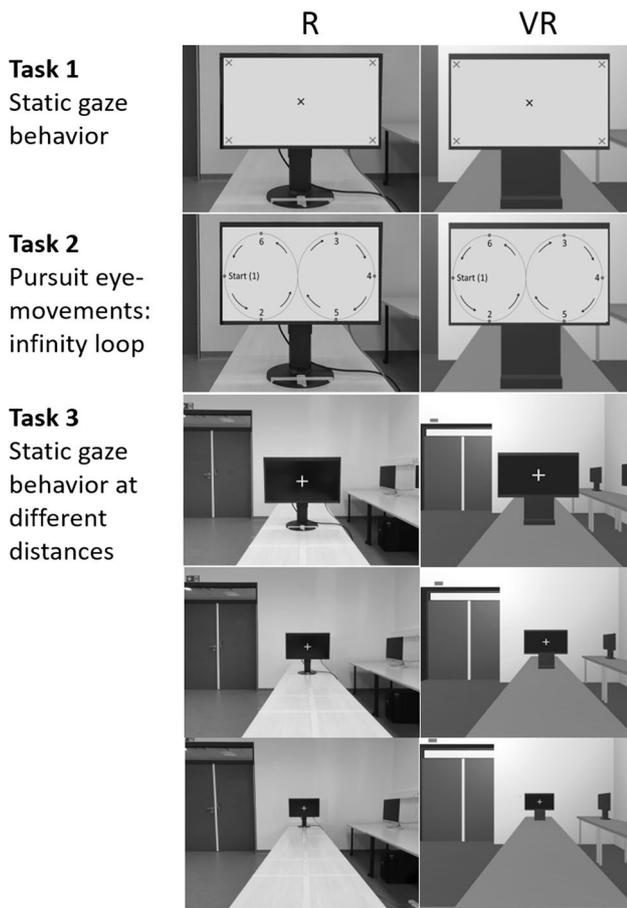


Fig. 3 Overview of all tasks from the participant's perspective. R: reality. VR: virtual reality

Unity 2018.3 (Unity Technologies, U.S.A.), and the data was accessible through the official plugin provided by SMI (iViewNG HMD Api Unity Wrapper v1.1, 2017).

3.3 Experimental protocol

The participants were randomly assigned into two groups: group 1 started the experiment in VR and group 2 began in R-condition. After installing the hardware components, the calibration was performed and the tasks were carried out in their predefined order (see Fig. 3). Each participant had to do each task twice. After completing each task per condition, the participant could relax and read the instructions for the next task. Subsequently, the participants changed the conditions (R/VR). After the participants completed all tasks, they were asked to complete a feedback questionnaire. The whole experiment took around 30 min per participant.

3.3.1 Preparation

In R, the height of the center of the monitor was adjusted to the participant's eye level in the seated position. The participant was fitted with the mobile eye tracker, which was firmly fixated on the participant's head. To gain reliably eye-tracking data from the HMD, it was important to adjust the individual interpupillary distance for each participant (Dörner et al. 2013). The HMD was placed on the head of the participants and they could then adjust the pupil distance themselves until they had a clear view. The participants were seated in front of a table, which was the same in R and VR ensuring equal haptic feedback in both conditions. The preparation was identical for each task.

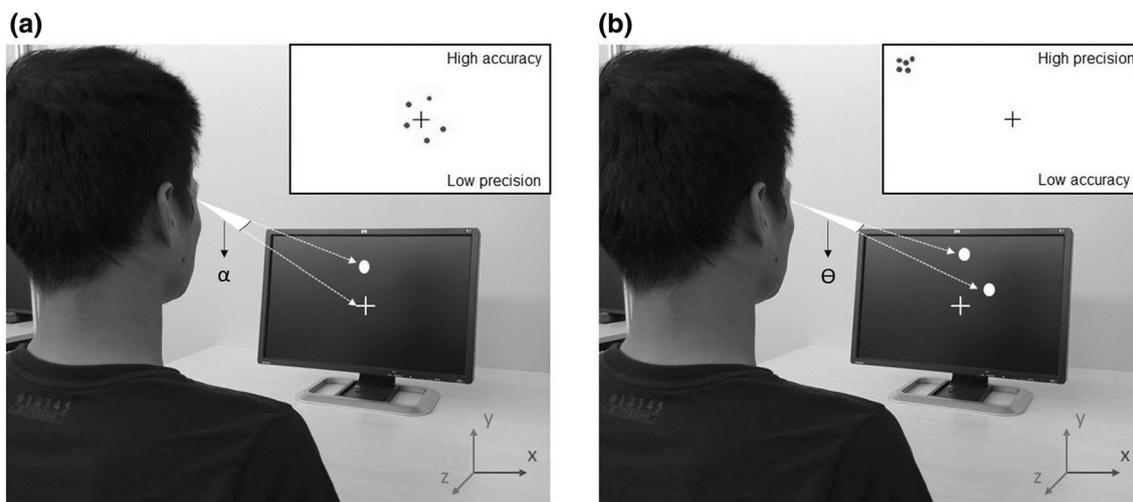


Fig. 4 Overview of gaze accuracy (a) and gaze precision (b) based on Nyström et al. (2013). The dots indicate the point of regard (POR) and the crosses indicate the target stimulus. The white boxes provide an example for a high accuracy but low precision and b high precision and low accuracy.

The arrows indicate the angle for each parameter. For better representation, exaggerated values for accuracy and precision have been used. Below on the right is the coordinate system, which was used for the calculation of the deviating angles

3.3.2 Calibration

A 3-point calibration was conducted for both devices according to the manufacturer's calibration protocol. The HMD was installed on the participant's head to be in the best position for the eye-tracking recommended by the manufacturer. Before each recording or trial, the calibration was repeated to avoid a loss of data quality over time due to a reduced shifting of the measurement system. The preparation and calibration procedures were identical in all three tasks.

3.4 Parameters

The algorithms calculating the accuracy and precision of the different measuring systems are based on raw data. In Fig. 4, both parameters are visualized. Gaze accuracy can be explained by the averaged distances between the position of the participants' gaze point and the target stimuli (gaze accuracy). The precision values indicate an averaged distance between each gaze point made by the participant. Accordingly, high accuracy and precision are characterized by low values. Meaning, the lower the angle between the two vectors (a) α for gaze accuracy and (b) θ for precision (see Fig. 4), the smaller the gaze deviation, and hence the higher the gaze accuracy and precision.

3.4.1 Accuracy (offset)

According to Holmqvist et al. (2015), the same formulas were used to calculate the average accuracy of the participants over the angle distribution (α) (see Fig. 4). The accuracy α_{Offset} results from the mean value, which corresponds to the recording frequency of the eye-tracking systems, and the mean value of all calculated angular deviations.

3.4.2 Precision (root mean square)

The same procedure or formula was used for the precision, instead, here the deviation of the distance was not determined from the reference cross (stimuli) to the (made) point of regard (POR), but the chronological sequence of the PORs recognized by the system. By using the root-mean-square (RMS), the quadratic mean was obtained, which in turn was calculated in deviation of the degree (Holmqvist et al 2015). The angle calculation is also used to extract the precision of the eye-tracking system. This is determined by the angle between two successive positions of the pupil cross (Holmqvist et al. 2015). θ represents the angle between the two vectors of each made gaze point (see Fig. 4). The squares of all angle values calculated in

a POR (of a cross) were summed up and divided by the number of data samples for the quadratic mean value.

3.4.3 Algorithm for accuracy and precision

To calculate the angles of accuracy and precision, we assume that the position of the participant's eyes was fixed in space, and the distance between the eyes and the monitor was constant. Through this, we could create an abstract coordinate system with the origin being the central point of the monitor. The X-axis was in the horizontal direction, the Y-axis was in the vertical direction, and the Z-axis pointed towards the participant (see Fig. 4). Therefore, the coordinate of the participant's eyes can be defined as point $P(0, 0, 100)$, because they were seated 100 cm in front of the screen. The next step was to calculate the angle between the eyes and the targets for the two parameters mentioned above. Therefore, we converted all the coordinates of the PORs and the relevant targets from pixels to centimeters. Then, the vectors from the eyes to the targets and the vectors from the eyes to the PORs were obtained. With these vectors, the dot product was used to calculate the angle accuracy using the formula shown below. The same idea was implemented for precision with a slight modification of the input vectors. For precision, the vectors from the eye position to each POR were calculated as well as the angles between each vector over time.

3.5 Data processing

To calculate the deviation of the system's registered gaze to the target (accuracy) and the deviations of the PORs among themselves (precision), the following steps were applied. When extracting the data, we were able to use the eye tracker's coordinate systems provided by the measuring systems. The origin of the 2D image of the two systems was defined at the top left corner. For each trial, the coordinates of the target stimuli as well as the point of regards (POR) were determined. The different pixels within the horizontal (X-axis) and vertical (Y-axis) direction were calculated by using the Pythagorean theorem to determine the size of the direction vector between them. Thus, we calculated the Euclidean distance from each POR to the target stimulus. All PORs were recorded and evaluated within an area of interest (AOI, in the form of a circle with a circumference of 3°). This ensured to avoid influences on gaze accuracy and precision by measuring those PORs, which were recorded between the reference cross and the target stimuli. To compare our results with results of other studies, the calculation of the deviations in angles for both, accuracy and precision, was conducted by using the previously mentioned vector calculation.

The statistical evaluation was performed with IBM SPSS Statistics 25. The algorithms for calculating the angular deviations were implemented in MATLAB 2018b (The MathWorks, U.S.A.). In total, the data sets of 21 participants were available for statistical analysis. The verification of significant differences was performed above an alpha level of 0.05. Pearson's correlations coefficient (r) were used to indicate the effect sizes.

3.6 Task description

3.6.1 Task 1: Static gaze behavior

3.6.1.1 Conduction Instead of using concentric circles as stimuli (Clemotte et al. 2014), we used crosses in the current study (see Fig. 3). A cross in the center of the screen was used as a reference to the other crosses. The other crosses appeared at the corners of the screen for 1.8 s. We wanted to record the gaze data for each cross for at least one second, so we added 0.8 s. The idea was to analyze whether participants were able to see fast emerging stimuli in VR. Each cross was displayed four times in a randomized order so that the participants could not predict where to look next. The reference cross remained visible at all times. The participants were instructed to fixate the middle cross as the new starting position after each fixation of one of the crosses at the corners was made. Each cross was presented for 7.2 s, bearing in mind that the reaction time must be subtracted from the participant's observation. The participants were asked to make as few blinks as possible when the stimulus targets appeared to ensure good data quality and to reduce problematic data collection.

3.6.1.2 Data analysis The univariate ANOVA with repeated measures for two paired samples and t test comparisons with calculated effect sizes were used to analyze differences in the gaze accuracy for each positioned cross [top right, bottom right, top left, bottom left]. For precision, a nonparametric Friedmann test of differences and Bonferroni-corrected post hoc comparisons with calculated effect sizes were conducted to analyze possible differences between each positioned cross.

3.7 Task 2: Pursuit eye-movements

3.7.1 Conduction

In this task, a dot appeared on the monitor (left side). This dot moved across the monitor in the form of an infinity loop for 15 s. The participants had to follow it with their eyes until the blue dot returned to the origin of the movement trajectory and stopped moving.

3.7.1.1 Data analysis A nonparametric Friedman test of differences and Bonferroni-corrected post hoc comparisons with calculated effect sizes were performed to analyze possible significant differences of gaze accuracy for each reference point (see Fig. 3). The center was not taken into account in the analysis, because gaze accuracy and precision were already examined in the other tasks.

3.8 Task 3: Static gaze behavior at different distances

3.8.1 Conduction

In the third task, a white cross appeared on a black background. Inside the white cross, a small black cross was visible so that the participants would not have any difficulty in discovering the center of the cross, especially for the 1 m distance. For each distance, the participant should fixate the cross for 3 s to ensure that they did not stare at the same target for too long and lose concentration in the process. After the fixation was finished, the monitor was set to the next distance (1 m, 2 m, and 3 m). Afterward, the monitor was repositioned and a new calibration was carried out.

3.8.1.1 Data analysis In the third task, a nonparametric Friedman test and Bonferroni-corrected post hoc comparisons with calculated effect sizes were also applied due to a lack of normal distribution. The gaze accuracy and precision for all distances [1 m, 2 m, 3 m] between both conditions [VR, R] were compared.

4 Results

4.1 Task 1: Static gaze behavior

Table 1 shows the results with no significant differences regarding gaze accuracy in task 1. It shows the basic level of information required in order to assess eye movement research (Holmqvist et al. 2012). In R, most participants fixated the top right (TR) and in VR the top left (TL) accurately concerning the different directions. For both measuring systems, the lowest accuracy was achieved at the low crosses, which was bottom left in R and the bottom right in VR.

The data across the different positioned stimuli were checked for normal distribution (Kolmogorov–Smirnov, $p = 0.200$). The Levene test showed equal variances ($p = 0.121$). A one-way ANOVA was conducted to compare the effect of the position of the cross (top left, top right, bottom left, bottom right) on the gaze accuracy (deg) between VR and R conditions. An analysis of variance showed no significant differences between the differently positioned crosses (top left, top right, bottom left, bottom right) in

Table 1 Comparison of the gaze accuracy and precision between R and VR for each positioned cross

Cross position	R <i>M</i> ± <i>SD</i> (deg)	VR <i>M</i> ± <i>SD</i> (deg)	<i>z</i> values, significance	Effect size (<i>r</i>)
<i>Accuracy</i>				
TR	0.46 ± 0.25	0.44 ± 0.23	<i>z</i> = 0.541, <i>p</i> = 0.595	No effect
BR	0.50 ± 0.27	0.62 ± 0.36	<i>z</i> = -1.494, <i>p</i> = 0.151	No effect
TL	0.59 ± 0.31	0.42 ± 0.30	<i>z</i> = 1.889, <i>p</i> = 0.073	No effect
BL	0.63 ± 0.36	0.55 ± 0.33	<i>z</i> = 1.209, <i>p</i> = 0.209	No effect
total	0.55 ± 0.30	0.51 ± 0.31	<i>z</i> = 1.054, <i>p</i> = 0.305	No effect
<i>C</i>	0.41 ± 0.08	0.39 ± 0.10	<i>z</i> = 0.713, <i>p</i> = .476	No effect
<i>Precision</i>				
TR	0.04 ± 0.02	0.06 ± 0.02	<i>z</i> = 3.190, <i>p</i> < 0.01	0.49
BR	0.04 ± 0.02	0.07 ± 0.02	<i>z</i> = 3.619, <i>p</i> < 0.01	0.59
TL	0.06 ± 0.04	0.07 ± 0.03	<i>z</i> = 1.286, <i>p</i> = 0.89	No effect
BL	0.05 ± 0.03	0.07 ± 0.03	<i>z</i> = 1.905, <i>p</i> = 0.12	0.29
Total	0.05 ± 0.02	0.07 ± 0.02	<i>z</i> = 2.520, <i>p</i> = 0.12	0.39
<i>C</i>	0.03 ± 0.01	0.07 ± 0.07	<i>z</i> = -3.980, <i>p</i> < 0.001	0.61

Precision values reflect the RMS of inter-sample distances. *M* mean, *SD* standard deviation, *TR* top right, *BR* bottom right, *TL* top left, *BL* bottom left and *C* center, *r* = Pearson's correlation coefficient indicates the effect size

VR and R with $F(3, 164) = 2.531$, $p = 0.059$. Based on the results of the ANOVA, relevant conditions were compared pair-wise by means of *t* tests, which revealed no significant differences between R and VR for each cross (see Table 1). The accuracy expressed by the degree of distribution in R and VR was around 0.5° ($R = 0.55^\circ$ and $VR = 0.51^\circ$).

The precision values (see Table 1) were also analyzed for possible statistical differences between the crosses in task 1. The data across the different positioned stimuli were checked for normal distribution (Kolmogorov–Smirnov, $p < 0.005$). A nonparametric Friedman test of differences was conducted and rendered a Chi-square value of 60.000, which was significant ($p < 0.001$). There is a difference in gaze precision between VR and R regarding the different positions of the crosses (see Table 1). Bonferroni-corrected post hoc comparisons indicated a significant difference in gaze precision with partly strong effect sizes between VR and R, except for the top left cross.

In addition, the difference (in degree) of gaze accuracy between the center and the corners of the screen was examined. A nonparametric Friedman test of differences was conducted and rendered a Chi-square value of 18.43, which was significant ($p < 0.001$). Bonferroni-corrected post hoc comparisons indicated that the mean score at the center of the screen in R ($M = 0.41$, $SD = 0.08$) was significantly lower than at the corners ($M = 0.55$, $SD = 0.30$). The same was observed in VR. The mean of the accuracy in the center ($M = 0.39$, $SD = 0.10$) was also significantly lower compared to the mean of the corners ($M = 0.51$, $SD = 0.31$). There was no significant difference between the center of R ($M = 0.41$, $SD = 0.08$) and VR ($M = 0.39$, $SD = 0.10$). No significant

difference in gaze accuracy in the corners between R ($M = 0.55$, $SD = 0.30$) and VR ($M = 0.51$, $SD = 0.31$) could be observed. In contrast to accuracy, the precision values differ between R and VR for the stimuli placed in the center and corners (except top left).

4.1.1 Discussion

The results show that there is no difference between R and VR within the directional vision measuring gaze accuracy. In VR, a better accuracy of 0.04° was obtained, which is not significantly different to R. Furthermore, for both conditions the highest accuracy was shown in the center of the field of view (FOV). The accuracy in the middle of the FOV, compared to the corners, was significantly better by 0.14° in R and by 0.12° in VR. The gaze accuracy ($R = 0.55^\circ$ and $VR = 0.51^\circ$) is in line with the manufacturer's specifications, which stated a gaze accuracy of 0.5° (iViewETG User Guide Version 2.7 2016). The fact that the accuracy at the center of the FOV is more accurate than at the corners, is also in line with the study of Hornof and Halverson (2002). However, this was not observed in the study of Nyström et al. (2013), in which different calibration methods were tested. Targets placed off-center did not differ in offset as compared to those positioned centrally. This previous result shows that the lowest accuracy was detected in the lower right corner (Feit et al. 2017). The most inaccurate measurement in R was in the lower-left corner. The gaze data of the HMD in task 1 are in line with those from the mobile Eye-Tracking system. Regarding the current data, it can be concluded that the visual information processing related to stimuli in a short

distance (a distance of 1 m), which are displayed in different directions, works similarly in VR compared to R. In the context of sports science, it is important to recognize a variety of visual stimuli and to react to them. The results of this task suggest that the operating mode of the visual system in VR can be carried out in the same way and the gaze behavior seems to be as accurate compared to R. Despite the significant differences within the precision values in the comparison between the realities, the quality of precision in VR is comparable with other devices and is precise enough to determine the parameters such as fixations and saccades.

4.2 Task 2: Pursuit eye-movements

To compare the data between VR and R of the infinity loop, the six fix points of the curve (Fig. 5) were selected.

The distance deviation was compared for each point in the infinity loop. A nonparametric Friedman test of differences among repeated measures was conducted and rendered a Chi-Square value of 118.38 which was significant ($p < 0.001$). Accordingly, the eye accuracy between the points differed. Bonferroni-corrected post hoc comparisons indicate a significant difference between each selected point (except for point 6) of the curve with strong effect sizes. The

deviation of degrees shows a significantly lower accuracy for VR with 2.76° (SD 0.86°) compared to R with 0.72° (SD 0.12°) (Table 2).

4.2.1 Discussion

In the pursuit eye-movement task, a highly significant difference was found between VR and R in eye-tracking movements within gaze accuracy. Six points were selected from the infinity loop to make further comparisons between pursuit eye-movements in R and VR. The points were determined by six specific points in time, which were selected manually before. To successfully implement such a task in the Unity Engine, and to be able to extract valid data afterward, a different design than a PowerPoint presentation as a video on an object (monitor) in the virtual scene should be chosen. A possible approach would be to implement an object (in our case a point) into the VR scene and let it migrate as an infinity loop as shown in task 2. This would generate access to the x and y coordinates and could determine the exact time of the maxima of the curve. One possible explanation for these differences (except for point 6) seems to be the significantly poorer resolution of the HMD. It may have been more difficult in VR to detect the visual

Fig. 5 The six fix points of the infinity loop. The arrows indicate the direction of movement

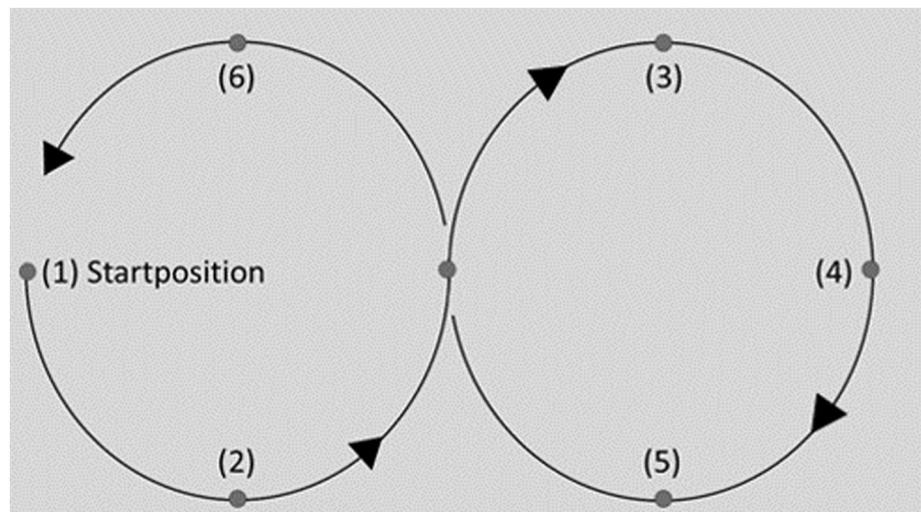


Table 2 Comparison between the gaze accuracy of R and VR by using the six points of the infinity loop as reference points (see Fig. 3)

Position	R <i>M</i> ± <i>SD</i> (deg)	VR <i>M</i> ± <i>SD</i> (deg)	<i>z</i> values, significance	Effect size (<i>r</i>)
1	0.53 ± 0.32	3.34 ± 2.08	$z = -6.725, p < 0.001$	1.04
2	0.68 ± 0.36	2.11 ± 0.68	$z = -5.450, p < 0.001$	0.84
3	0.92 ± 0.55	4.05 ± 0.55	$z = -4.500, p < 0.001$	0.69
4	0.52 ± 0.36	1.70 ± 0.57	$z = -5.325, p < 0.001$	0.82
5	0.92 ± 0.46	2.59 ± 0.56	$z = -4.275, p < 0.001$	0.66
6	0.73 ± 0.59	1.35 ± 0.86	$z = 0.525, p = 0.645$	No effect

M mean, *SD* standard deviation, *r* = Pearson’s correlation coefficient indicates the effect size

stimuli compared to R. This also emerged from the questionnaires of the participants, who experienced difficulties in perceiving the moving point in some places. In addition, the center of a circle may be more difficult to fixate than a center of a cross. This is, of course, a critical factor, especially concerning gaze accuracy. Authors emphasized the difficulty in distinguishing between system errors and a non-existent view of the target (Dalrymple et al. 2018). By using highly developed head-mounted displays, however, this factor could be limited. For faster movements, the authors suggested the use of devices with a higher measurement frequency, whereas 60 Hz is described as a too low frequency (Gibaldi et al. 2017). The frame rates of the different applications could differ (Clay et al. 2019). Although the current 3D scene was created without any complex computations, the quality of gaze measurements could have suffered, due to a limited synchronization and frame interpolation between the lower frame rate of the game engine (Unity) and the eye-tracker (Clay et al. 2019). Since gaze accuracy can also be influenced by calibration (Nyström et al. 2013), it should be mentioned that both devices (SMI mobile Eye Tracker and HMD integrated Eye-Tracker System) are based on a three-point calibration method that is system-controlled in VR and operator-controlled in R. Unfortunately, in our case, it was not possible to change the calibration method manually for the SMI devices. The operator-controlled calibration was shown to be preferred over the system-controlled calibration, which is considered the worst of all (Holmqvist et al. 2012). In general, they found that when participants were allowed to perform the calibration themselves, the accuracy and precision of the gaze data were significantly the best. Since no differences were found in the first task, the different calibration methods may affect the accuracy of a stimulus that moves continuously more severely than a static stimuli in VR. This could be verified by testing devices against each other by using the same calibration method while examining moving targets. Another reason could be the lack of

experience of the participants within the VR. Only six of them had previous experiences but did not have their own VR glasses for private use, which suggests that their experience was relatively low. The results of the questionnaire show that one of them needed a break or complained about cybersickness. Nevertheless, to pursue a moving stimulus seems to be a challenge for the current used VR application. Therefore, a similarity of gaze accuracy between VR and R must be falsified. According to the results from the current study, the accuracy of the visual system in VR is much worse for dynamic stimuli and should be considered during the development of moving visual cues.

4.3 Task 3: Static gaze behavior at different distances

4.3.1 Between-condition comparison

A nonparametric Friedman test of differences was conducted and rendered a Chi-square value of 37.952, which was significant ($p < 0.001$). There is a difference in gaze accuracy over the different positioned stimuli (see Table 3). Bonferroni-corrected post hoc comparisons indicated that there is no significant difference between VR and R for the 1 m condition. The Wilcoxon test shows that there was a medium-strong significant difference between VR and R from the measurement of the deviating distance overall in gaze accuracy. The precision data revealed no normal distribution (Kolmogorov–Smirnov, $p < 0.001$). A nonparametric Friedman test of differences was conducted and rendered a Chi-square value of 67.449, which was significant ($p < 0.001$). In this task, there was also a difference in gaze precision between VR and R. Bonferroni-corrected post hoc comparisons indicated a difference for the 1 m condition ($z = 2.905$, $p < 0.001$, effect size $r = 0.45$), for the 2 m condition ($z = 2.333$, $p < 0.001$, effect size $r = 0.36$), and for the 3 m condition ($z = 2.810$, $p < 0.001$, effect size $r = 0.43$).

Table 3 Comparison between the gaze accuracy and precision of R and VR of the different distances between the conditions (R vs. VR)

Distance (m)	R $M \pm SD$ (deg)	VR $M \pm SD$ (deg)	z values, significance	Effect size (r)
<i>Accuracy</i>				
1	0.41 ± 0.08	0.39 ± 0.10	$z = 0.571$, $p = 0.322$	No effect
2	0.21 ± 0.00	0.38 ± 0.12	$z = -1.190$, $p = 0.39$	0.18
3	0.17 ± 0.04	0.38 ± 0.09	$z = -2.667$, $p < 0.001$	0.41
Total	0.27 ± 0.15	0.39 ± 0.21	$z = -2.068$, $p = 0.039$	0.32
<i>Precision</i>				
1	0.03 ± 0.01	0.07 ± 0.07	$z = -2.905$, $p < 0.001$	0.45
2	0.03 ± 0.01	0.07 ± 0.12	$z = -2.333$, $p < 0.001$	0.36
3	0.03 ± 0.01	0.08 ± 0.09	$z = -2.810$, $p < 0.001$	0.43
Total	0.03 ± 0.00	0.07 ± 0.09	$z = -4.015$, $p < 0.001$	0.62

M mean, SD standard deviation, r = Pearson's correlation coefficient indicates the effect size

Regarding the comparisons between R and VR, there is no significant difference at the 1 m distance. A significant difference was observed for the 2 m distance, but only with a small effect size. The difference is more obvious within the 3 m distance where a large effect was detected. Compared to the results of task 1 (see Table 1), gaze accuracy seems to be at its best level in the center of the screen in both conditions (R and VR). The quality of gaze accuracy is influenced by the position of the presented stimuli, as it decreases when fixating at larger eccentricities. These results are an additional factor that proves the similarity of gaze accuracy in both systems (mobile Eye-Tracker in R and Eye-Tracker in HMD, both SMI). For precision, no significant differences could be found between the center and the corners of the screen ($p > 0.05$).

4.3.2 Within-condition comparison

The differences over the distances can be explained through the different characteristics of the continuity of each measurement system (see Table 4). In VR, the system works constantly regarding the gaze accuracy over the three fixation crosses at different distances. Therefore, no statistical difference between 1 m, 2 m, and 3 m in VR-condition was observed (all $p > 0.05$). It turned out that the course of the accuracy differed within both measuring systems. In VR, the accuracy remained at the same level over the three distances. When comparing the distances among themselves,

Table 4 Comparison between the gaze accuracy and precision of R and VR of the different distances within each condition, r =Pearson's correlation coefficient indicates the effect size

Condition	Distance	z values, significance	Effect size (r)
<i>Accuracy</i>			
R	1 m versus 2 m	$z = 1.952, p = 0.001$	0.30
	2 m versus 3 m	$z = 0.952, p = 0.099$	No effect
	3 m versus 1 m	$z = 2.905, p < 0.001$	0.45
VR	1 m versus 2 m	$z = 0.190, p = 0.741$	No effect
	2 m versus 3 m	$z = 0.524, p = 0.364$	No effect
	3 m versus 1 m	$z = 0.333, p = 0.564$	No effect
<i>Precision</i>			
R	1 m versus 2 m	$z = 0.190, p = 0.741$	No effect
	2 m versus 3 m	$z = 0.429, p = 0.458$	No effect
	3 m versus 1 m	$z = 0.619, p = 0.284$	No effect
VR	1 m versus 2 m	$z = 0.667, p = 0.248$	No effect
	2 m versus 3 m	$z = 0.143, p = 0.805$	No effect
	3 m versus 1 m	$z = 0.524, p = 0.364$	No effect

no significant difference was detected (all $p > 0.05$). By focusing on the R-condition, we found differences between 1 and 2 m and between 1 and 3 m. No significant difference between the 2 m and 3 m distance was detected. The differences between R and VR can be explained by the continuous improvement in gaze accuracy over further distances with the mobile eye-tracker (R). For further detail, see Table 4.

Similar to the accuracy, the precision values in R decrease with increasing distance, while they remain relatively constant in VR. However, in contrast to accuracy, the changes within each condition are not significant ($p > 0.05$).

4.3.2.1 Discussion In the third task, no differences in gaze accuracy were found between VR and R at the 1 m distance, similar to task 1. Within 2 m distance, there is a significantly lower accuracy in R compared to VR, but with a small effect size. Only from a distance of 3 m, a large effect was observed. These differences increase if the pixels are not adjusted over the distances in R. Accordingly, it must be taken into account that the number of pixels of a 2D image is distributed differently in size to different distanced objects in the scene. In VR, the coordinate system of the game engine (Unity3D) can be used and the relations between pixel and real distance are calculated automatically. While the accuracy of the R-condition improves with increasing distance, it remains constant in VR (see Table 3). The results are not surprising. The lower screen resolution in VR compared to R could lead to difficulties in perceiving the center of the fixation cross. In R, they still could perceive the center, whereas in VR, they often reported focusing just at the white fixation cross which reveals no accurate observation. Nevertheless, the deviation of the fixations produced by the participants from the target stimulus in VR is only around an angle deviation of 0.39° , which reveals a sufficient ability to observe other people or objects in daily situations or more specifically opponents, teammates movements', or sport equipment motions in sports scenarios. The different deviations of the two measuring systems might be affected by the different quality of stimuli presentation. The accuracy of the visual system can also be described as sufficient in this task. Compared to the fixation crosses performed in the current study, the stimuli from the sport science context (ball, bat, opponent, teammate, body regions, etc.) are much larger and therefore easier to recognize in VR.

The precision values differ between R and VR for all distances with strong effects (see Table 3). Nevertheless, the precision values of the integrated eye-tracker in the HMD are still comparable to those from other measurement systems mentioned by Nyström et al. (2013). This allows the detections of fixations and enables a comparison between individuals during participants' activities or sports performances in VR. Nevertheless, when observing the standard deviation (SD) of the precision values (see Table 3), abnormally high

values could be detected. The HMD rendered two images for both eyes at the same time to create a stereo view in the VR. However, there seems to be a dark area in the middle of FOV, which blocked the real content in the scene when the user stared at this area. In this task, the cross was placed right in the middle of the screen for observation. When this cross was rendered for each eye, its position in the FOV was very close to this blocked area and this may explain the large SD value in precision because the participant was trying to find the cross in the middle (see Fig. 3). In the first task, however, these high values within the SD were not observed (see Table 1). This leads to the assumption that the discrepancy is not only due to the stimuli placed at different positions in task 3 but that it is also an issue due to different kinds of stimuli presented in each task (see Fig. 3).

5 General discussion

In the current study, the accuracy and precision of the visual system were measured and compared between the real and virtual conditions. Different stimuli were used to confront the visual system in various ways. The static crosses were placed at the corners of the screen and in the center. In addition, the participants had to permanently observe a point moving across the screen presented as an infinity loop. Furthermore, the participants had to look at static crosses in the center of the monitor. By modifying the monitor's position in relation to the participant, the fixations took place at different distances. The three tasks were chosen because an easily feasible implementation of the study in VR could take place. Due to the reference cross in the center of the screen, it was easy to calculate the length of the gaze vectors as well as the distance between the position of the target and participants' gaze point. These should be the first step to compare gaze accuracy and precision between R and VR. Perceiving stimuli placed on different positions at the monitor is an often-used method to calculate participants' gaze accuracy and precision (Feit et al. 2017; Hornof and Halverson 2002; Holmqvist et al. 2012). Since the manufacturer of the mobile eye-tracker and the integrated eye-tracker is the same in the HMD, a better insight into the behavior of participants was attempted to reach. The assumption that each device is equipped with the same technique and uses the same algorithms allows the conclusion of possible differences of the two systems due to foveal gaze behavior of the participants.

The within-subject design allows a direct comparison of VR and R. Each participant underwent the VR and R scene, which reduced the possibility of finding differences in the results due to different eye physiologies, varying neurology, and psychology, different ability to follow instructions, wearing glasses or contact lenses or having long eyelashes or droopy eyelids, which all can influence the quality of the

data (Nyström et al. 2013). The homogeneity of the testing participants was ensured. Therefore, sports students at the same age and pedigree were chosen for participation in the current study. Participants wearing glasses were rejected due to problems with the installation of the different hardware systems simultaneously. However, no official test design for eye quality was conducted, which could be helpful to exclude possible outliers. Also, Nyström et al. (2013) stated the operator's experiences could affect the data quality. For each R and VR, only one operator was involved in the conduction to reduce possible influences. Both of them were well instructed and had to go through several test runs before starting the experiment.

Throughout the results of task 1, it can be said that the gaze accuracy within VR coincides with that of reality. Therefore, the greatest similarity occurred at a distance of 1 m. Although there is a difference between the two measurement systems at further distances, the calculated accuracy in the VR is still sufficient to ensure that the participants consider the implemented stimuli in the experiment. Since the lower resolution made the perception of the stimuli more difficult, it is conceivable that these differences would no longer occur with a higher resolution.

Looking at the 1 m distance, the accuracy does not differ between R and VR. Nevertheless, even if the values differ at different distances between the conditions, there is no concern with static stimuli. Previous studies have shown that values from 0.7° to 1.3° are found to be an acceptable indicator for SMI applications (Blignaut 2009). For dynamic stimuli presented in task 2, critical values that are above this defined threshold were found. When considering the precision of all tasks, the values are similar to those of the studies carried out so far, even if other systems were used (Holmqvist et al. 2011). To verify the accuracy and precision only from the influence of the measuring system, it is recommended to use an artificial eye, as it does not generate any movements of its own, where precision values of 0.001° – 1.03° were observed (Holmqvist et al. 2011). To be able to make further conclusions here, more data needs to be generated with other HMDs since the data between tower-mounted and remote devices already differ. Furthermore, the limitation that the calibration method brings with it should be discussed. In VR, a system-controlled calibration method was used, which in any case is valid for the lowest precision value compared to the other methods (Nyström et al. 2013). An implementation of a self-executable calibration method of the integrated eye-tracking system in HMD could increase the accuracy to set the fixated positions at the right time. If a technical implementation was provided, an examination with an artificial eye would be helpful to test the true values of both precision and accuracy (Nyström et al. 2013). Poor precision can be determined by the quality of the eye camera and the algorithms that determine the position of the

pupil and the corneal reflection. The only constant that the VR environment creates is the lighting condition. To check this more in detail, other VR devices that have an integrated eye tracker should be tested. Higher developed devices with higher resolution should be integrated since lower resolution could lead to lower precision values. The eye-tracker in VR records data with a lower frequency than the mobile one that was used for R. However, recording samples with 30 Hz compared to those of 60 Hz have no impact on precision values (Ooms et al. 2015).

In the current study, the participants were placed in a seated position in front of a monitor and the visual stimuli were presented via an integrated PowerPoint presentation. This setup does not provide any information about gaze behavior to a highly dynamic situation in real-world sports. Kredel et al. (2017) suggest testing the sports-related perceptual-cognitive skills under more realistic conditions. Nowadays, mobile eye-trackers can be connected to smartphones or portable laptops, which ensures the recording of eye movements during the performance of more complex or extensive movements compared to those of the current study. The mobility of the integrated eye-tracking system in the HMD is restricted due to the length of the cable and is therefore difficult to use during real-world sports scenarios, as it occurs in most of the HMD eye trackers (Clay et al. 2019). An additional technical factor could be latencies that can occur due to the representation of the VR scenario of the game engine through the cable-based HMD. Further investigations have to be done to reveal those durations. Normally, the mobile eye-tracking system provides an external camera that shows the gaze pattern of the point of regards distributed over the FOV. Predefined Areas of interest (AOI) were used for analyzing the gaze pattern. This function can also be implemented in the VR since AOIs can be freely chosen via the integration of objects getting hit by a gaze vector. In addition, it is possible to trace the time when previously defined regions were looked at (Clay et al. 2019). Therefore, a comparison between both conditions can be ensured. Nevertheless, this study does not comply with all quality criteria for sports-related eye-tracking research because of the missing realistic viewing condition or naturalistic response like real movements (Kredel et al. 2017). Further studies need to be conducted, which have to meet these criteria and at the same time measure the gaze behavior of the participants conducting sports activities. Fast head movements could be the most challenging factor that needs to be solved.

In sports, it is also fundamental to perceive other objects in the environment without fixating them by using the peripheral vision (Vater et al. 2017). The current study is focusing on the foveal vision. Since peripheral vision plays an important role in the decision to act, this should also be investigated in a further study in VR. The results show that as long as the visual stimuli are well represented in the VR

environment and a smooth perception can occur, studies can also be carried out with this application. The current results endorse the use of gaze data in VR. As long as the headset is mounted correctly on the head and the additional time for the re-calibration procedure can be accepted, Eye-Tracking in VR is considered to be a useful tool (Clay et al. 2019).

6 Conclusions

The results of the present study show that regarding the 1 m distance and static visual stimuli equal gaze accuracy between VR and reality could be observed. For precision, a worse result has been detected. Despite the difference to the mobile eye tracker, the integrated system in the HMD works precisely enough, as can be seen from the small deviations of the precision values. The results reflect realistic values so that an analysis of gaze behavior can also be carried out in the VR with reliable and valid data for fundamental research. We assume that an improved resolution of the HMD and the presentation of easily recognizable stimuli lead to accurate and precise fixations. When stimuli can be detected without any difficulties by the user, VR can be a valuable possibility to create sport relevant training scenarios with individualized visual cues. Based on the results of the current study, dynamic stimuli were perceived worse than static ones. Moreover, the stimuli moved in a predictable path, which does not correspond to realistic conditions. For this purpose, further data with improved measurement techniques and extended stimulus presentation (size, predictable and unpredictable trajectories, distance, characteristics, etc.) must be collected. During the conduction of the current study, the distances of the experimental setup always remained the same in VR, since they were fixed in the scene and therefore allow accurate measurements and controlled conditions. The current VR applications can already create a good feeling of immersion that encourages the user to act as in reality. However, to do more than just fundamental research, the technical components need to be improved. Further investigations in a more realistic scene have to be done to provide suggestions on how visual training in VR should look like for improving athletes' performances.

Acknowledgements Open Access funding provided by Projekt DEAL. The project was financed by the German Research Foundation (DFG) under Grant WI 1456/22–1.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in

the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Blignaut P (2009) Fixation identification: the optimum threshold for a dispersion algorithm. *Atten Percept Psychophys* 71(4):881–895
- Blignaut P, Wium D (2013) Eye-tracking data quality as affected by ethnicity and experimental design. *Behav Res Methods* 46(1):67–80
- Clay V, König P, König SU (2019) Eye tracking in virtual reality. *J Eye Mov Res* 12(1):3. <https://doi.org/10.16910/jemr.12.1.3>
- Clemotte A, Velasco M, Torricelli D, Raya R, Ceres R (2014) Accuracy and precision of the Tobii X2-30 Eye-tracking under non ideal conditions. In: Londral AR (ed) Proceedings of the 2nd international congress on neurotechnology, electronics and informatics, Rome, Italy, 25–26 October, 2014, pp 111–116. S.I.: SCITE-PRESS. <https://doi.org/10.5220/0005094201110116>
- Dalrymple KA, Manner MD, Harmelink KA, Teska EP, Alison JT (2018) An examination of recording accuracy and precision from eye tracking data from toddlerhood to adulthood. *Front Psychol* 9:803. <https://doi.org/10.3389/fpsyg.2018.00803>
- Drewes J, Montagnini A, Masson GS (2011) Effects of pupil size on recorded gaze position: a live comparison of two eyetracking systems. *J Vis* 11(11):494
- Dörner R, Broll W, Grimm P, Jung B (2013) Virtual and augmented reality (VR/AR). Basics and methods of virtual and augmented reality. Springer, Berlin, p 48
- Duque G, Boersma D, Loza-Diaz G (2013) Effects of balance training using a virtual-reality system in older fallers. *Clin Interv Aging* 8:257–263
- Düking P, Holmberg H-C, Sperlich B (2018) The potential usefulness of virtual reality systems for athletes: a short SWOT analysis. *Front Physiol* 9:128. <https://doi.org/10.3389/fphys.2018.00128>
- Feit AM, Williams S, Toledo A, Paradiso A, Kulkarni H, Kane S, Morris MR (2017) Toward everyday gaze input: accuracy and precision of eye tracking and implications for design. In: Mark G, Fussell S, Lampe C, Schraefel MC, Hourcade JP, Appert C, Wigdor D (eds) Proceedings of the 2017 CHI conference on human factors in computing systems—CHI '17. ACM Press, New York, pp 1118–1130. <https://doi.org/10.1145/3025453.3025599>
- Gibaldi A, Vanegas M, Bex PJ, Maiello G (2017) Evaluation of the Tobii EyeX Eye tracking controller and Matlab toolkit for research. *Behav Res Methods* 49(3):923–946
- Gray R (2017) Transfer of training from virtual to real baseball training. *Front Psychol* 8:2183. <https://doi.org/10.3389/fpsyg.2017.02183>
- Holmqvist K, Nyström N, Andersson R, Dewhurst R, Jarodzka H, van de Weijer J (2011) Eye tracking: a comprehensive guide to methods and measures. Oxford University Press, Oxford
- Holmqvist K, Nyström M, Mulvey F (2012) Data quality: what it is and how to measure it. In: Proceedings of the 2012 symposium on Eye-tracking research and applications. ACM, pp 45–52
- Holmqvist K, Nyström M, Andersson R, Dewhurst R, Jarodzka H, van de Weijer J (eds) (2015) Eye tracking: a comprehensive guide to methods and measures (first published in paperback). Oxford University Press, Oxford
- Hooge ITC, Holleman GA, Haukes NC, Hessels RS (2018) Gaze tracking accuracy in humans: one eye is sometimes better than two. *Behav Res Methods*. <https://doi.org/10.3758/s13428-0181135-3>
- Hornof AJ, Halverson T (2002) Cleaning up systematic error in eye-tracking data by using required fixation locations. *Behav Res Method Instrum Comput* 34(4):592–604
- Kredel R, Vater C, Klostermann A, Hossner E-J (2017) Eye-tracking technology and the dynamics of natural gaze behavior in sports: a systematic review of 40 years of research. *Front Psychol* 8:1845. <https://doi.org/10.3389/fpsyg.2017.01845>
- Krokos E, Plaisant C, Varshney A (2019) Virtual memory palaces: immersion aids recall. *Virtual Real* 23:1–15. <https://doi.org/10.1007/s10055-018-0346-3>
- Loomis J, Knapp J (2003) Visual perception of egocentric distance in real and virtual environments. In: Hettlinger J, Haas MW (eds) Virtual and adaptive environments: applications, implications, and human performance issues. Lawrence Erlbaum Associates Publishers, pp 21–46. <https://doi.org/10.1201/9781410608888.pt1>
- Messing R, Durgin FH (2005) Distance perception and the visual horizon in head-mounted displays. *ACM Trans Appl Percept* 2:234–250. <https://doi.org/10.1145/1077399.1077403>
- Molina KI, Ricci NA, de Moraes SA, Perracini MR (2014) Virtual reality using games for improving physical functioning in older adults: a systematic review. *J NeuroEng Rehabil* 11(1): Article no. 156
- Neumann DL, Moffitt RL, Thomas PR, Loveday K, Watling DP, Lombard CL, Antonova S, Tremee MA (2018) A systematic review of the application of interactive virtual reality to sport. *Virtual Real* 22(3):183–198. <https://doi.org/10.1007/s10055-017-0320-5>
- Nyström M, Holmqvist K (2010) An adaptive algorithm for fixation, saccade, and glissade detection in eye-tracking data. *Behav Res Methods* 42(1):188–204
- Nyström M, Andersson R, Holmqvist K, van de Weijer J (2013) The influence of calibration method and eye physiology on eyetracking data quality. *Behav Res Methods* 45(1):272–288
- Ooms K, Dupont L, Lapon L, Popelka S (2015) Accuracy and precision of fixation locations recorded with the low-cost Eye Tribe Tracker in different experimental setups. *J Eye Mov Res* 8(1):5
- Petri K, Bandow N, Witte K (2018a) Using several types of virtual characters in sports—a literature survey. *Int J Comput Sci Sport* 17(1):1–48. <https://doi.org/10.2478/ijcss-2018-0001>
- Petri K, Ohl C-D, Danneberg M, Emmermacher P, Masik S, Witte K (2018b) Towards the usage of virtual reality for training in sports. *Biomed J Sci Tech Res* 7(1):1–3. <https://doi.org/10.26717/BJSTR.2018.07.001453>
- Petri K, Emmermacher P, Danneberg M, Masik S, Eckardt F, Weichelt S, Bandow N, Witte K (2019) Training using virtual reality improves response behavior in karate kumite. *Sports Eng* 22:2. <https://doi.org/10.1007/s12283-019-0299-0>
- Pfeuffer K, Mayer B, Mardanbegi D, Gellersen H (2017) Gaze + Pinch interaction in virtual reality. In: Proceedings of the 5th symposium on spatial user interaction (SUI '17). ACM, New York, pp 99–108. <https://doi.org/10.1145/3131277.3132180>
- Reichert E (2019) Genauigkeit und Präzision des Eye-Tracking Systems—Dikablis Professional unter standardisierten Bedingungen (Masterarbeit)
- Renner RS, Velichkovsky BM, Helmer JR (2013) The perception of egocentric distances in virtual environments—a review. *ACM Comput Surv* 46:1–40. <https://doi.org/10.1145/2543581.2543590>
- Rose FD, Brooks BM, Rizzo AA (2005) Virtual reality in brain damage rehabilitation: review. *Cyberpsychol Behav* 8(3):241–262
- SensoMotoric Instruments (2009) iView X system manual (version 2.4) [Computer software manual]. Berlin, Germany.
- SensoMotoric Instruments (2016) iViewETG user guide (version 2.7) [Computer software manual]. Berlin, Germany
- Tirp J, Steingröver C, Wattie N, Baker J, Schorer J (2015) Virtual realities as optimal learning environments in sport—a transfer

study of virtual and real dart throwing. *Psychol Test Assess Model* 57(1):57–69

Vater C, Kredel R, Hossner E-J (2017) Examining the functionality of peripheral vision: from fundamental understandings to applied sport science. *Curr Issues Sport Sci*. https://doi.org/10.15203/CIS_2017.010

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.