2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Hess Screen Revised: How Eye Tracking and Virtual Reality change Strabismus Assessment

Wolfgang Mehringer[1], Markus Wirth[1], Franka Risch[1], Daniel Roth[4], Georg Michelson[2,3] and Bjoern Eskofier[1]

Abstract— Strabismus is a visual disorder characterized by eye misalignment. The extent of ocular misalignment is denoted as the deviation angle. With the advent of Virtual Reality (VR) Head-Mounted-Displays (HMD) and eye tracking technology, new possibilities measuring strabismus arise. Major research addresses the novel field of VR strabismus assessment by replicating prism cover tests while there is a paucity of research on screen tests. In this work the Hess Screen Test was implemented in VR using a HMD with eye tracking for an objective measurement of the deviation angle. In a study, the functionality was tested and compared with a 2D monitor-based test. The results showed significant differences in the measured deviation angle between the methods. This can be attributed to the type of dissociation of the eyes.

Clinical relevance— HMDs offer a high degree of dissociation and a consistent measurement environment. The advantage of eye tracking is the low level of user interaction required, which makes it accessible to almost all patients.

I. INTRODUCTION

Strabismus describes the imbalance of the muscles responsible for the eye movements, resulting in eye misalignment. Approximately 70% of the population show a latent (hidden) deviation which can be treated by corrective visual aids [1], [2]. However, a manifest misalignment (prevalence of 4%) is usually treated by surgery [3]. The prerequisite for successful treatment is a detailed measurement and evaluation of the present deviation angle [1], [2].

Established methods are light reflex tests in which the corneal reflections indicate the deviation, cover tests in which compensatory eye movements are used to estimate the deviation and screen tests in which two objects have to be visually superimposed by the subject. However, these methods are partly inaccurate due to being dependent on the examiners' level of experience as well as complex to implement [4]. The Hess Screen Test [5] allows a detailed insight into the ocular deviation pattern in nine directions of sight while the cover tests only assess strabismus at central gaze. The Hess Screen Test shows targets at 0° and 15° deviation to the center position. The targets are visible to the fixating eye, while the pointer with which the patient is to superimpose the targets is visible to the tested eye [6]. Gaps between target and cursor quantify the deviation angle.

Further disadvantages of this test are its high degree of user interaction, the dependence on external influences such as lighting conditions as well as the inability to measure latent deviations in some patients. Thus, several alternative approaches were developed using eye tracking, partly also in combination with Virtual Reality (VR) Head-Mounted-Displays (HMD). Eye tracking reduces the need of user interactions since it directly measures where the user is looking at. The advantage using HMDs is the high-degree dissociation of the eyes, but also a consistent measurement environment can be established (e.g. lighting conditions).

This work aims to combine the advantages of eye tracking and VR on the basis of the Hess Screen Test to enable a detailed measurement and evaluation of the deviation angle. The Hess Screen Test for this work was implemented as 2D monitor-based test to obtain reference measurements and as novel VR version, that measures the deviation angle with a minimum of user interaction solely using eye tracking. Within a study, 12 healthy subjects were tested with both methods. The measured deviations are then compared.

II. BACKGROUND & RELATED WORK

The Hess Screen test allows conclusions on cause and extent of the deviation. The examiner points to a position on the chart using a green laser while the patient tries to superimpose it with a red one. Anaglyph glasses dissociate the eyes so that the patient only sees one laser on each eye [7]. If strabismus is present the lasers cannot be superimposed and the resulting gap quantifies the deviation. However, screen tests require the patient to cooperate, the dissociation by anaglyph glasses is not optimal (i.e. due to environmental influences) so that patients still see both lasers on one eye and the results depend on the examiner's experience [8].

To compensate for the subjective assessment by the examiner Thomson et al. developed a computerised version of the Hess Screen Test using a 2D monitor screen and the computer mouse for pointing [9]. However, the ocular dissociation still relies on complementary colour filters which may lead to poor dissociation of the eyes.

supported by the FEDERAL MINISTRY FOR ECONOMIC AFFAIRS AND ENERGY (BMWi)

[1]Machine Learning and Data Analytics Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Germany correspondence to wolfgang.mehringer@fau.de

[2]Department of Ophthalmology, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Germany

[3]Talkingeyes&More GmbH, Erlangen, Germany

[4]Human-Centered Computing and Extended Reality, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Germany

Awadein et al. compared the results of 82 patients using two computerised screen tests [10]. The evaluation showed great agreement between both computerised versions and the traditional test. Thorisdottir et al. came to similar conclusions comparing a digital screen test with the traditional one [8].

To minimize the patient's interactions, Bakker et al. [11] developed a system consisting of three cameras and a set of infrared LEDs that record the eyes. Their results show a standard deviation of the error angle less than 0.5°. These findings are in line with comparable research work also utilizing eye tracking [12]–[15] with screen or cover tests.

To gain independence from environmental influences and good dissociation of the eyes, Nesaratnam et al. applied VR technology [16]. They compared a traditional screen test with a VR-based test. The overlapping task was performed with a computer mouse. Results showed a good agreement between the two versions. Their findings are supported by Mehringer et al. [17] who developed an overlapping task using the VR controllers as well as Miao et al. who focused on the cover-uncover and alternating prism cover test using a VR headset with an integrated eye tracking system to display the targets and record the eye movements [18].

Previous research shows that computerised tests provide equivalent results to their conventional counterparts. It is also shown that eye tracking can improve the test by minimizing the user's interaction making results more objective and that VR can be applied to minimize undesired confounding factors within the environment. However, to the best of our knowledge, there is no VR system using eye tracking that performs a screen test like the Hess Screen Test. Therefore, the aim of this work is to evaluate the applicability of the Hess Screen Test utilizing a controlled VR environment in combination with eye tracking.

III. METHODS

A. VR-based Measurement Environment

The VR setup was implemented using Unreal Engine 4.22.3 and SRanipal for eye tracking functionalities. Figure 1 shows the structure of the scene. It includes a grid in the background providing one highlighted focus point as the target (i.e. center position). The targets and the user are in a static relationship to each other, meaning that the view onto the targets does not change with head movements by the user. The grid is located in $x = 50\,cm$ distance in x-direction to the user and the targets are positioned at $\alpha, \beta \in \{-15°, 0°, 15°\}$, with α referring to the horizontal or y-direction, and β to the vertical deviation or z-direction, like in the traditional Hess Screen Test. The angles are converted to coordinates, so that all target positions $\vec{t}$ can be described by the following vector relative to the user's location:

$$\vec{t} = x\left(1, \frac{\tan(\alpha)}{\cos(\beta)}, \frac{\tan(\beta)}{\cos(\alpha)}\right) \quad (1)$$

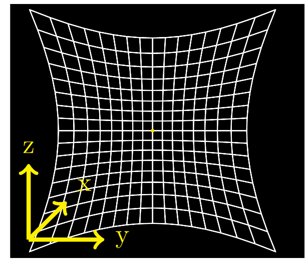

Fig. 1. Structure of the scene with the target at center position $\alpha = \beta = 0°$. The lines in the figure are thickened for better visualization. The actual grid in VR view is thinner. Each intersection correlates to 5°.

For the target fixation a focus vector is defined as the difference of the gaze origin $\vec{g_O}$, and the gaze direction $\vec{g_D}$ of the specified eye $\vec{f} = \vec{g_D} - \vec{g_O}$. If the highlighted target is within a 3 cm radius around the focus vector, a hit is detected. We tested the software and proofed that a direct hit is not feasible as the accuracy of the eye tracker and microsaccades by the user must be taken into account. On the first target hit, a time frame of two seconds is triggered in which the focus vectors of following successful hits are registered for both eyes. To distinguish between latent and manifest strabismus we created two stages that differ only in the representation of the grid and the targets.

1) Simulation Stage I: During the test, the targets for the nine directions of sight are shown two times and in a random order to each eye. Furthermore, the visibility of each new target and the background grid switches between the eyes so that both eyes can be tested simultaneously. In that way targets and the grid are only visible to one eye at a time. This results in 36 focus targets. If the target is visible to the left eye the behavior of the left eye fixating the target is assessed as well as the behavior of the right eye and vice versa.

2) Simulation Stage II: To test if the measured deviation from stage I indicates latent or manifest strabismus, the contained objects of stage I (i.e. grid and highlighted target) are made visible to both eyes simultaneously. The timer and thus a new target can be triggered independently by both eyes. However, the target must be hit by at least one of the two eyes. The number of targets is reduced to 18 by no longer differentiating between the two eyes. If a deviation occurs in stage I but not in stage II a latent strabismus is present. A deviation in both stages indicates a manifest strabismus.

3) Data Processing: The measured focus vectors of stage I and II were grouped according to the target's visibility to left or right eye. The measurements were plotted on typical Hess charts showing the recorded measurements like shown in Figure 2. Since the gaze behavior of both eyes is recorded, the deviation angle can be calculated as the Euclidean distance between the median focus points of each eye for each target respectively.

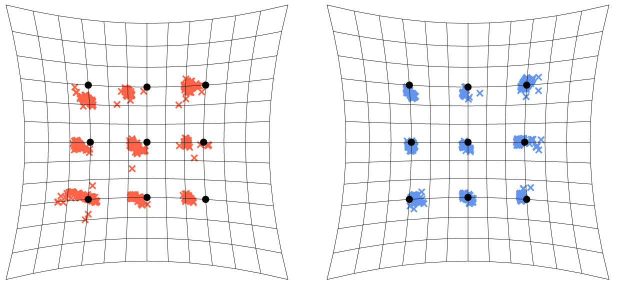

Fig. 2. Representation of the visual evaluation of the deviation measurements. This example depicts all focus points of both eyes measured with the target only visible to the right eye.

B. 2D Monitor-based Measurement Environment

For reasons of comparison, a 2D monitor-based version of the Hess Screen Test was designed using Python 3.8. The anaglyph glasses were equipped with colour filters in red and blue. The subject is presented red dots as targets in nine directions of sight equivalent to stage I. In contrast to the VR test, there is no grid in the background, since this would affect the dissociation of the eyes. A chin rest ensures that the subject's head is centered and at a distance of 50 cm from the screen. The subject must hit each new target with a blue cursor, controlled with a computer mouse. By left-clicking the mouse a new target is triggered. Each target is displayed once. The deviation angle was calculated from the Euclidean distance between the target position and the measurement.

C. Study Design

A within-subjects design repeated-measure study was conducted at the hospital of the Friedrich-Alexander-University Erlangen-Nuremberg (FAU) including 12 healthy subjects. On average, the subjects were 25 years old. The study was approved by the FAU Ethics Committee and conducted following the Declaration of Helsinki. The university's hygiene rules were applied to reduce risk of spreading Sars-CoV-2.

After giving consent to the study and providing demographic data the subject's visual acuity was estimated with the Freiburg Visual Acuity Test[1] (FrACT), starting with the dominant eye which was determined using an ocular dominance test. Both of the screen tests were performed without auxiliary means like glasses. To avoid sequencing effects, the order of tests was counterbalanced between the subjects. The interpupillary distance (ipd) of each subject was assessed using a tape measure and applied to the lenses of the HMD HTC Vive Pro Eye (HTC Corporation, Taoyuan, Taiwan) with integrated eye tracker that was used for the study. After successfully calibrating the eye tracker and headset for the VR test the Hess chart was displayed to both eyes with the nine targets visible to familiarize the subjects with the task. For the monitor-based test, the utilized anaglyph glasses ensured a blue filter in front of the

[1]https://michaelbach.de/fract/index.html

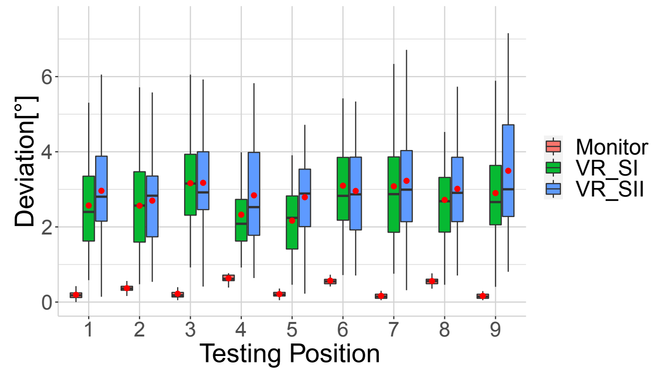

Fig. 3. Boxplots including all considered deviations, ordered for method and testing position. VR_SI stands for the results from the VR-based method stage I while VR_SII indicate the results from stage II. The red dots indicate the mean values.

tested eye. Next, the subject was presented all targets including the background grid to check equal visibility. After the deviation of the first eye was measured for all nine directions, the same procedure applied for the other eye respectively. Each subject completed each test three times for more data points.

The calculated median ocular deviations derived from the VR-based and the monitor-based method were tested for normal distribution with the Shapiro-Wilk test. Since the Shapiro-Wilk tests showed predominantly no normal distribution ($p < 0.05$) pairwise Bonferroni corrected Wilcoxon signed rank tests were performed based on the testing eye and grouped by the testing position. The tests were performed for each of the methods individually to assess the effect of the testing eye. Since the tests indicate no significant difference between the testing eyes for both monitor-based ($p > 0.1$) and VR-based method ($p > 0.1$) the variable was excluded from further analysis. Thus, we computed two Friedman tests in unreplicated complete block design with either the testing method or the testing position as block variable . Since the Friedman showed an effect based on the different testing methods, the data was further aggregated and again tested with Friedman (subject ID as blocking variable). For each Friedman test the Kendall's coefficient of concordance Wt was evaluated as effect size. Next pairwise Bonferroni corrected Wilcoxon signed rank tests with Cohen's d as effect size were performed to compare the different testing methods for significance. Bonferroni was applied to reduce the chance of type I error in significance testing. Wilcoxon signed rank tests and Friedman tests were applied as non-parametric test due to results on the Shapiro Wilk test. We expect small to no deviation in the monitor-based version (inability to measure latent deviations) while stage I of the VR-based method should result in higher deviations (latent deviation in 70% of population [1], [2]). In stage II of the VR-based method the found deviations of stage I should disappear.

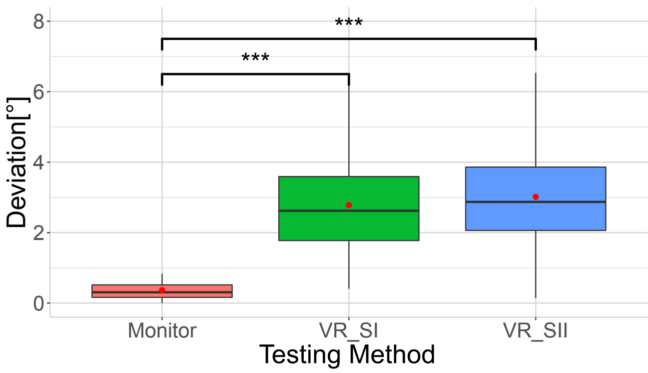

Fig. 4. Boxplots including the different testing methods and the resulting significances: *** $p <= .001$. The red dots indicate the mean values.

IV. RESULTS

In Figure 3 the measured deviations are shown according to measurement method and testing position. The Euclidean distances calculated from the measured values of the monitor-based ($Monitor$) test are low compared to the VR-based methods (VR_SI and VR_SII). The shown diagram excludes outlier for better visualization. The results of the Friedman test with the testing method as block variable revealed no significant difference across the testing positions ($X^2 = 9.244, p > 0.3$). The Kendall's W for the same configuration indicates a medium to small effect size ($WT = 0.385$). Further, the Friedman test with the testing position as block showed a significant difference across the testing methods($X^2 = 14.889, p < 0.001$) and a large effect size ($WT = 0.827$). Based on the result of the two Friedman tests the data was aggregated based on the testing positions and again tested with the subject ID as blocking variable. The results again show a significant difference across the testing methods with a large effect size ($X^2 = 55.389, p < 0.01, WT = 0.769$). For pairwise comparison of the testing methods Bonferroni corrected Wilcoxon signed rank tests with Cohen's d as effect size were performed resulting in significant differences between Monitor and VR_SI ($p < 0.01, |d| > 1$) as well as Monitor and VR_SII ($p < 0.01, |d| > 1$). There is no significant difference between VR_SI and VR_SII ($p > 0.5, |d| < 0.1$). The results are summarized in Figure 4.

V. DISCUSSION

As expected the results of our conducted study show significantly higher measured deviations in the VR-based methods compared to the monitor-based one. This may be due to differing dissociation methods of the eyes. Since the VR headset provides completely separate images to the eyes this system has a higher degree of dissociation compared to complementary colour dissociation [6] making latent deviations visible [4]. Conversely, the red-blue colour dissociation was not sufficient to do so.

Eight of the considered subjects (66%) showed a deviation in stage I and II that was not visible in the monitor-based one. Considering that mild heterophoria is common for the vast majority of the people in close sight (approximately 70% [1], [2]) this finding is reasonable, however, the found deviation in stage II is not. A reason for this could be measurement errors of the eye tracking due to readjustments of the headset. Since the eye tracking did not always immediately recognise a target hit, most subjects had to adjust the VR-goggles to continue the test leading to a corrupt eye calibration.

The eye tracker only specifies an accuracy of 0.5°–1.1° within a FoV of 20°. This means that targets with a deviation of 15° may already have higher measurement inaccuracies. Furthermore, the median focus vectors used to calculate the deviation may also disguise the true underlying deviation and hence lead to a systematic error of the system.

VI. CONCLUSIONS

In this paper, a new version of the Hess Screen Test, was developed that combines the advantages of objective measurement using eye tracking with the environmental consistency offered by VR HMDs.

To reduce measurement errors of the eye tracking, it might be useful to implement an individual calibration procedure. This may result in more reliable detection of target hits and less need for readjustment. It would also be interesting to test whether targets with only 10° offset lead to more precision in using eye tracking. A study including patients with diagnosed strabismus could provide more insight into the validity of the measurements.

The combination of objective measurement with the consistency of a VR environment is rich in opportunities for the application in ophthalmology. Especially the good dissociation and the low level of user interaction are advantageous.

ACKNOWLEDGMENT

Bjoern Eskofier gratefully acknowledges the support of the German Research Foundation (DFG) within the framework of the Heisenberg professorship programme (grant number ES 434/8-1).

References

[1] M. E. Razavi et al., “Normative values for the fusional amplitudes and the prevalence of heterophoria in adults,” Iranian Journal of Ophthalmology, vol. 22, pp. 41–46, 2010.

[2] M. Radakovic et al., “Heterophoria and fusional convergence and divergence in preschool children,” Medicinski Glasnik, vol. 9, pp. 293–298, 2012.

[3] H. Hashemi et al., “Global and regional prevalence of strabismus: a comprehensive systematic review and meta-analysis,” Strabismus, vol. 27, no. 2, pp. 54–65, 2019.

[4] K. W. Wright et al., Handbook of Pediatric Strabismus and Amblyopia. Springer, 2006.

[5] W. R. Hess, “Ein einfaches messendes Verfahren zur Motilitatspruefung der Augen,” Zeitschrift fuer Augenheilkunde, vol. 15, pp. 201–219, 1916.

[6] J. M. Roodhooft, "Screen Tests Used to Map out Ocular Deviations," Bull. Soc. belge Ophtalmol, vol. 305, pp. 57–67, 2007.

[7] G. K. Von Noorden and E. C. Campos, Binocular Vision and Ocular Motility: Theory and Management of Strabismus. Mosby, 2002.

[8] R. L. Thorisdottir et al., "Comparison of a new digital KM screen test with conventional Hess and Lees screen tests in the mapping of ocular deviations," Journal of American Association for Pediatric Ophthalmology and Strabismus, vol. 22, pp. 277–280, 2018.

[9] W. D. Thomson et al., "A new system for the measurement of ocular motility using a personal computer," Ophthalmic Physiological Optics, vol. 8, 1990.

[10] A. Awadein, "A computerized version of the lancaster red-green test," Journal of American Association for Pediatric Ophthalmology and Strabismus, vol. 17, pp. 197–202, 2013.

[11] N. M. Bakker et al., "Accurate Gaze Direction Measurements With Free Head Movement for Strabismus Angle Estimation," IEEE Transactions on Biomedical Engineering, vol. 60, pp. 3028–3035, 2013.

[12] K. P. Weber et al., "Strabismus measurements with novel video goggles," American Academy of Ophthalmology, vol. 124, pp. 1849–1856, 2017.

[13] Z. H. Chen et al., "Eye-tracking-aided digital system for strabismus diagnosis," Healthcare Technology Letters, vol. 5, pp. 1–6, 2018.

[14] O. Yehezkel et al., "Automated diagnosis and measurement of strabismus in children," American Journal of Ophthalmology, vol. 213, pp. 226–234, 2020.

[15] Y. Zheng et al., "Intelligent evaluation of strabismus in videos based on an automated cover test," Applied Sciences (Switzerland), vol. 9, no. 4, 2019.

[16] N. Nesaratnam et al., "Stepping into the virtual unknown: Feasibility study of a virtual reality-based test of ocular misalignment," Macmillan Publishers Limited, part of Springer Nature, vol. 31, pp. 1503–1506, 2017.

[17] W. A. Mehringer et al., "An image-based method for measuring strabismus in virtual reality," in 2020 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), 2020, pp. 5–12.

[18] Y. Miao et al., "Virtual reality-based measurement of ocular deviation in strabismus," Computer Methods and Programs in Biomedicine, vol. 185, pp. 1–9, 2019.