Evaluaton of the low budget do it yourself eye tracking system: Your Eye Tracker (YET)

Florian R. Bender

Master Thesis

Faculty of Behavioural, Management and Social Sciences (BMS),

University of Twente

Human Factors and Engineering Psychology

First Supervisor: Martin Schmettow

Second Supervisor: Cesco Willemse

2024

Abstract

In this thesis, we aim to investigate the potential of YET (Your Eye Tracker) as an ultra-low-budget do-it-yourself solution to make eye tracking technology more available and affordable. Our main objective is to systematically evaluate the accuracy and precision of YET compared to existing eye tracking systems. By conducting calibration tests and comparing data, we seek to determine the feasibility of YET for various applications. Employing a within-subject design, participants completed a calibration task before viewing stimuli displayed on both a white background (congruent with the calibration background) and a black background (incongruent with the calibration background). This variation aimed to examine the influence of screen background consistency on calibration accuracy. Additionally, three head-mounted lighting conditions (LED, near-infrared, and ambient lighting only) were tested to assess their effects on YET's performance under varying illumination. Results indicate that YET is able to achieve an accuracy close to established eye tracking systems at the center of the screen but shows reduced accuracy toward the outer edges of the screen. Incongruency of the screen´s background to the calibration screen resulted in a noticeable misalignment between stimuli and estimated gaze points whereas alignment was consistent in the congruent background condition. Variations in lighting conditions (LED, NIR, or no additional light) did not appear to impact YET’s performance.

Introduction

Vision is often considered to be our most crucial and intricate sense, playing a fundamental role in how we perceive, interpret, and interact with the world. The eyes have been called the window to the soul, a metaphor that underscores the significance of vision in understanding not only the external environment but also the nuances of human behaviour and intent. As our dominant sensory modality, vision enables us to interpret subtle facial expressions, particularly through the eyes, and to convey complex social cues, such as intent, emotion, and understanding, in our interactions. The study of visual perception and eye movements thus provides a unique lens through which we can analyze both physiological and psychological processes.

In recent decades, eye tracking technology has emerged as a powerful tool for investigating human cognition, attention, and social interaction. By capturing and analyzing eye movements, researchers can gain insights into oculomotor, cognitive, and neurobiological processes (Holmqvist & Andersson, 2017). According to Holmqvist and Andersson (2017), recording eye movements can be beneficial for advancing scientific knowledge and driving technological progress.

Despite its proven value, however, eye tracking technology remains underutilized and often underrepresented in social science education, where opportunities to engage with specialized instruments are more limited compared to the biological sciences. For instance, students in biological sciences are frequently trained in the use of tools like microscopes, which support their experimental learning and skill development. In contrast, students in social science programs are less likely to receive comparable exposure to advanced research technologies, including eye tracking systems. This disparity is primarily due to the high cost and logistical requirements associated with eye tracking equipment, which many institutions cannot afford to integrate into their curricula yet (Rakhmatulin, 2020).

The lack of access to eye tracking technology in educational settings represents a missed opportunity for students to engage with emerging methodologies that can expand their research capabilities and methodological rigor. By learning to use and understand eye tracking systems, social science students could deepen their understanding of human behaviour, improve data collection techniques, and ultimately contribute to more nuanced and robust research findings. Moreover, (early) exposure to eye tracking technology could encourage students to explore innovative research avenues, potentially fostering advances in fields centered on human attention, decision-making, and behaviour.

Given the current barriers to access, there is an increasing call for more affordable and accessible solutions that could democratize the use of eye tracking technology across disciplines.

Addressing these challenges would not only benefit students in educational settings but also advance research efforts by equipping a broader range of scholars and practitioners with the tools needed to engage in eye-tracking-based studies. The development of low-cost, attainable alternatives could therefore play a transformative role in expanding the educational and research possibilities associated with eye tracking technology.

Brief history of eye tracking

Eye tracking technology has a rich history that dates back to the late 19th century when early experiments were conducted to understand eye movements and visual perception. The study of eye movements can be traced back to the work of French ophthalmologist Louis Émile Javal in the late 19th century. Javal conducted experiments to measure eye movements during reading and other visual tasks using simple observation techniques, such as a mirror placed to the side of one eye.

In the early 20th century, researchers such as Alfred L. Yarbus and Edmund Huey made significant contributions to the understanding of eye movements and visual attention through observational studies and experimental methodologies. Huey built a device that allowed him to identify which words people were focusing on (Housholder et al., 2021; Walczyk et al., 2014). In his experiments he observed that, while reading, participants did not have smooth eye movements moving horizontally along sentences but appeared to focus on some words while moving quickly through others (Housholder et al., 2021; Walczyk et al., 2014). He was among the first to show that reading involves saccadic jumps (Walczyk et al., 2014). Huey´s device functioned similarly to a seismograph (Höffner, 2018). Participants had a cup with a hole in the middle fitted on their cornea and the cup was connected to an aluminum stylus drawing the eye movements on a rotating drum (Walczyk et al., 2014). Similarly, Yarbus developed suction devices which were attached to the eyes (Tatler et al., 2010). Huey´s, Yarbus´ and other early measurement techniques fostered progression for the development of eye tracking but caused considerable discomfort for patients (Haslwanter & Clarke, 2010; Tatler et al., 2010).

Later advancements in recording techniques in the 20th century, such as electrooculography (EOG) and scleral search coils, allowed researchers to study various aspects of eye movements more accurately and precisely. EOG has been widely used since the mid-20th century and relies on the fact that the eye's cornea and retina have different electrical potentials (Haslwanter & Clarke, 2010). When the eye moves, the electric field across the eye changes, resulting in a measurable voltage difference. By placing electrodes strategically around the eyes, these voltage changes can be captured and analyzed to infer the direction and extent of ocular movements (Haslwanter & Clarke, 2010). The most common electrode configuration involves placing electrodes above and below one eye, creating a bipolar recording setup (Belkhiria et al., 2022).

Sclera search coils have been coined to be the “gold standard” for measuring eye movements (Houben, Goumans, & van der Steen, 2006; Shelhamer & Roberts, 2010). In 1963, the original device, made use of rigid contact lenses secured to the eyes using suction and connected to wire coils (Shelhamer & Roberts, 2010). Improvements were made by embedding the wire coils in a softer silicone base, making the lenses more comfortable and eliminating the need for suction (Shelhamer & Roberts, 2010). This updated version remains in use to this day (Shelhamer & Roberts, 2010). While sclera search coils are renowned for their combination of resolution, accuracy, low noise and fast response, the drawbacks include the fragility of the wires as they might cause discomfort, are prone to breakage leading to data loss and experiment disruption (Shelhamer & Roberts, 2010).
These recording techniques paved the way for the emergence of computer-based eye tracking systems.

The development of computer technology in the latter half of the 20th century facilitated the transition to computer-based eye tracking. Early computer-based eye trackers, such as the MIT I-SCAN system developed in the 1970s, used video cameras and image processing algorithms to track eye movements. Commercial eye tracking systems began to appear in the 1980s and 1990s, offering researchers and practitioners access to more sophisticated and user-friendly tools for studying eye movements.

Over the past few decades, eye tracking technology has seen significant advancements in both hardware and software capabilities. High-speed cameras, infrared illumination, and advanced algorithms have enabled more accurate and precise tracking of eye movements. Recent trends in eye tracking technology include the development of wearable eye trackers, remote eye tracking systems, and mobile eye tracking solutions, expanding the range of applications and research possibilities.

The history of eye tracking technology reflects a trajectory of continuous innovation and advancement, driven by the quest to understand the complexities of human vision and behaviour. From its early roots in observational studies to modern computer-based systems, eye tracking has evolved into a powerful tool with far-reaching applications and implications across various disciplines.

Video-based eye trackers

Video-based eye tracking represents a significant advancement over earlier, more invasive methods of tracking eye movement. Unlike mechanical trackers, video-based devices utilize cameras and image processing software to non-invasively capture the user’s gaze patterns (Hansen & Ji, 2009). This section explores the primary components, functionalities, and practical considerations of video-based eye trackers.

The primary hardware component in video-based eye tracking devices is the camera, which captures video footage of the user's eyes. Video-based eye trackers rely on one or more cameras to capture images of the eye and make use of image processing hardware to calculate the point of gaze (Hansen & Ji, 2009). These cameras can vary in type and quality, ranging from standard webcams or built-in computer cameras to specialized high-resolution or infrared (IR) cameras designed specifically for eye tracking. IR cameras are especially beneficial in low-light conditions, as they enhance eye visibility through infrared illumination, enabling more precise tracking and are not visible to the user. For instance, devices like Apple's Vision Pro Mixed-Reality Headset utilize IR illuminators, enhancing accuracy even in challenging lighting conditions.

Video-based eye trackers typically include specific hardware, such as mounting options. Depending on the application, these devices may either be screen-based (remote) systems, which do not physically contact the user, or head-mounted devices, like glasses or headsets, designed to move with the user`s head (Narcizo et al., 2013). Mounting hardware plays a key role in stabilizing the device and minimizing interference from head movements, a common issue in accurate eye tracking. Further, eye tracking devices often include headrests.

The software in video-based eye tracking devices includes image processing algorithms that analyze the video footage captured by the camera. These algorithms are responsible for detecting and tracking various features of the user's eyes, such as the pupil, iris, eye corners, and eyelids (Hansen & Ji, 2009). Eye detection and tracking algorithms identify the position, size, and orientation of the user's eyes within each frame of the video. For this, either shape- or appearance-based models can be used (Hansen & Ji, 2009). Shape-based models aim to remodel the iris as a circular object while appearance-based models discern between image features, such as contrast between the pupil and sclera (Hansen & Ji, 2009).

Software-based calibration and validation procedures are used to calibrate the eye tracking system. Calibration involves presenting visual stimuli at known locations on the screen and recording the corresponding eye movements. Validation verifies the accuracy of the calibrated eye tracker by comparing predicted gaze positions with observed data. The software interface of video-based eye tracking devices provides users with controls for initiating recordings, adjusting settings, performing calibrations, and viewing eye tracking data in real-time or post-processing. Data visualization tools may be included to display eye movement trajectories, heatmaps, fixation patterns, and other visualizations of gaze behaviour.

While video-based eye tracking technology offers versatility, several limitations can impact its accuracy. Potential hindrances and limitations might arise in the detection of the eye due to its large variability in the appearance and dynamics (Hansen & Ji, 2009). Additionally, factors such as changes in lighting conditions and interference through head movements or glasses can affect eye detection. Individual differences among users, such as variations in eye shape, size, and color, can pose challenges. Furthermore, user fatigue or discomfort during prolonged use of eye tracking systems can affect the quality of data collected. Users may experience eyestrain or tiredness, leading to changes in their eye movement patterns or behaviour.

In terms of practical considerations, the cost of eye tracking technology can be prohibitive for individuals or organizations. The high cost of hardware, software, and associated equipment, as well as the need for specialized expertise for setup and maintenance, can limit the accessibility of eye tracking technology for research fields, industries, and individuals. Eye tracking devices can cost up to tens of thousands of Euros (Rakhmatulin, 2020). Rakhmatulin (2020) reviewed low-cost eye-tracking devices ranging from 100€ to 1840€, which further substantiates the unaffordability of ´low-cost´ eye tracking as they still pose financial challenges for users or organizations, particularly in developing countries.

Janthanasub and Meesad (2015) compared three low-cost eye tracking devices (“GazePoint’s GP3”, “EyeTribe”, and “DIY eye tracker”) while testing them as potential source for computer input to aid communication for handicapped users. Their results indicate that the three devices, ranging from 70$ to 900$, can all be a potential valuable resource for human computer interaction research (Janthanasub & Meesad, 2015).

In contrast, initiatives such as from the IT University of Copenhagen, which developed an eye tracker with production costs as low as 30€, demonstrate a significant step towards affordable eye tracking technology (Mantiuk et al., 2012). Despite its low production cost of around 30 euros, it has been reported to have comparable precision to commercial devices (Mantiuk et al., 2012). Mantiuk et al. (2012) compared this low-cost head-mounted eye tracker of the IT University of Copenhagen to a high-end eye tracker and found that their accuracy is comparable at around 1°, after removing 10% of gaze point outliers. Since there is no standardized way to measure the accuracy of eye tracking devices yet, they used the calibration procedure as delivered by the manufacturer and measured the accuracy based on the raw and unfiltered gaze data while considering factors such as viewing angle, human traits and visual fatigue (Mantiuk et al., 2012). After a training session, participants were instructed to look at a set of 25 target points arranged in a 5x5 grid, each being displayed for 2 seconds in a random order (Mantiuk et al., 2012). The components of the eye tracker in use, called “DIY”, are said to cost under 30 euros, though in their study they used a patient support frame of an ophthalmic slit lamp (generally used for eye disease examination) that includes an adjustable chin rest and forehead band (Mantiuk et al., 2012).

YET, Your Eye Tracker, (https://github.com/schmettow/YET) takes this accessibility a step further by providing an ultra-low budget DIY option. By utilizing affordable components such as a cheap USB endoscope camera and a small 3D-printed part, YET empowers users to assemble their eye tracker, thereby fostering understanding for the technology behind eye tracking.

Through initiatives like YET, the barriers to entry for eye tracking technology are further lowered, enabling a wider range of individuals and organizations to explore its applications in research, education, and beyond. By promoting affordability and accessibility, these initiatives contribute to the progress of eye tracking technology, ultimately enriching its utility and impact across various fields and communities.

YET

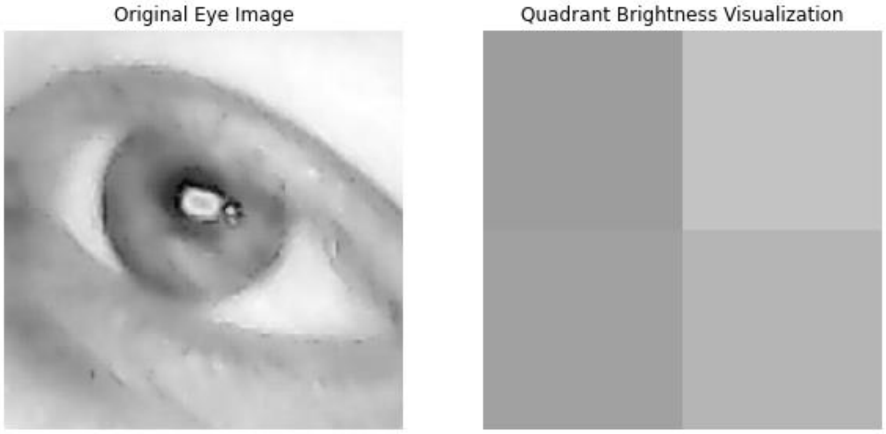
YET, which stands for "Your Eye Tracker," is a project aimed at making eye trackers affordable and attainable to everyone, particularly aimed at students in the social sciences. YET employs easily available and inexpensive materials. The core hardware components include an endoscope camera, a plastic ruler, glue, a tiny 3D printed part, a head mount (like a glass frame or helmet) and head rest (such as a water bottle, books or paper towels). These components are chosen for their affordability, ensuring that individuals with limited resources can still participate in building and using eye tracking devices.

Regarding the software, Yet interface (Yeti) is an open-source code. Yetis are software mechanisms responsible for processing the video stream captured by the endoscope camera into useful data, such as eye coordinates. Yetis utilize libraries like OpenCV for video processing and Pygame for user interaction. The different iterations of Yetis span from basic functionality, like simply displaying the video stream (as seen in Yeti0), to more advanced features such as implementing calibration points and statistical learning for accurate eye tracking (as seen in Yeti14).

YET makes use of a quadrant brightness-based model combined with multiple linear regression to estimate the participant’s eye movements. YET divides the image of the eye into quadrants and measures the relative brightness levels in each quadrant (Figure 1).

Figure 1

Eye image (left) and example of visualization of quadrant brightness (right)



With the use of a 3x3 calibration grid, the user is instructed to look at calibration stimuli. This process is applied to record quadrant brightness for corresponding target positions. The calibration data is then used to train a linear regression model to predict eye positions based on quadrant brightness patterns. By analyzing these brightness variations YET can estimate the position of the pupil. This data is then processed using multiple linear regression, which allows the software to predict gaze direction based on the relationships between quadrant brightness and corresponding eye coordinates.

Throughout the presentation of stimuli, the user is presented with a recalibration point at the center of the screen. By observing the eye's position relative to this known point, the system can recalibrate and adjust the eye tracking model.

Potential Applications of Affordable Eye Tracking

Eye tracking has found application in research of cognition, learning and social behaviour. As well as in the medical field for example for the diagnosis of eye related diseases (Li et al., 2021). Further use has been found in the commercial field for advertising research, product and web design, efficiency at work, sports, competitive games or other fields (Li et al., 2021). Customer analytics, recruitement, military research specifically in operations, instructions, UX & UR, VR and gaming all have made use of eye tracking systems to further their development. A more available eye tracking system might bear the potential to extend research in social sciences across the world and offer cross-cultural comparisons particularly with developing countries as well as promoting rep

Da Silva Lima and Ventura (2023) propose that eye tracking is particularly valuable in studies involving neurological patients or infants who may struggle to follow verbal instructions, grasp task requirements, or provide manual responses while serving the complementary function to aid in determining the conditions under which a response is given or offering the potential analysis of new dependent variables. Again, particularly in crisis areas of developing countries where communication might be a hindrance, an affordable eye tracking system could benefit professional care takers, patients and benefit progress in research.

To further understand the potential of YET it has to be tested and compared to existing eye tracking systems to determine its potential use of application.
Current eye trackers have the capabilities to measure and record the position and movement of the eyes. With the data, saccades, fixations, time to focus and duration of interest and revisits on areas of interest (AOI) can be defined. The benchmarks for evaluating eye trackers often include accuracy, precision, temporal resolution, and spatial resolution.

While studies involving saccade detection require eye trackers with a higher sampling rate, measured in Hz, studies focusing on particularly attention on areas of interest can be estimated to require less accuracy and precision especially when the areas of interest are larger. Saccade detection is particularly important in understanding how individuals move their eyes when shifting focus between different stimuli, as it provides insights into cognitive processes such as attention and information processing.

Saccade detection may play a role in various cases, such as in ophthalmology clinics were subtle abnormalities can be detected to diagnosing, monitoring or understand conditions such as glaucoma, macular degeneration and diabetic retinopathy (Tahri Sqalli et al., 2023). Contrary, in the diagnosis of ADHD and ASD patterns of visual attention are studied rather than saccade detection.

Levantini et al. (2020) reported and compared data of studies involving eye-tracking technology as a non-invasive methodology for testing clinical characteristics of ADHD, such as attention and inhibition control. The reported accuracy of the eye trackers that were used in the studies range from 0.5° to 2° of visual angle error; a lower value of accuracy represents less visual angle error.

Some of the reported studies examined saccades, which would require an eye tracker with a high sampling rate, while other studies compared the attention on areas of interest where accuracy and precision were less critical for the task. Comparing the time that people with ADHD focus on salient features of the face, distractors in pictures or even differing attention between whole pictures themselves presents the possibility that an eye tracker with less affordances to its accuracy might be able to yield similar results in various areas of use. Other areas, such as in education or for cognitive or commercial studies could make use of an affordable and available eye tracker, even with a worse accuracy, especially for attention studies differentiating between larger areas of interest.

Specifications of Eye tracking devices

Visual degree

Vision is measured in visual degrees. The visual field of humans comprises 120° horizontally and 90° vertically (Mantiuk, 2017), most of which is peripheral vision at above 30° in diameter. For angles higher than 20° it is natural to change the head position (Mantiuk, 2017). The visual angle of one eye can be calculated from the point of the center of the eye. The formula for calculating the visual angle (α) is:

$$\alpha = \frac{180}{\pi} \times 2 \times \arctan\left(\frac{h}{2d}\right)$$

Where h is the size of the object and d the distance from the eye to the object.

Accuracy

Accuracy in eye tracking refers to the distance θ_i (in visual degrees) between the actual x_i and the measured $\hat{x}_i$ gaze position, whereas the actual gaze position, or fixation refers to the point at which a person is expected to look at (Holmqvist, Nyström, & Mulvey, 2012; Reingold, 2014). According to Holmqvist & Andersson (2017) there is little consensus on how a fixation should be defined or measured since we cannot know where a human is looking (Holmqvist, Nyström, & Mulvey, 2012). The accuracy, given n recorded samples, can be calculated as followed:

$$\theta_i = \sum_{i=1}^{n} \frac{180}{\pi} * 2 * \arctan\left(\frac{x_i - \hat{x}_i}{2d}\right)$$

Precision

Precision in eye tracking refers to the consistency or repeatability of gaze measurements (Holmqvist & Andersson, 2017). (Spatial) precision refers to the variance in accuracy (Holmqvist, Nyström, & Mulvey, 2012). It quantifies how closely repeated measurements of gaze positions cluster around a central point. Mathematically, precision can be calculated using the standard deviation (SD) of the recorded gaze positions. The formula for calculating precision using SD is:

$$\sigma = \sqrt{\sum_{i=1}^{n} \frac{180}{\pi} * 2 * arctan\left(\frac{x_i - \hat{x}_i}{2d}\right)^{2}}$$

Brand et al. (2021) note that accuracy and precision should not solely be evaluated in isolation. A participant´s data recording might be accurate and precise, precise but not accurate, accurate but not precise and neither precise nor accurate (Brand et al., 2021). Between participants and screen location this interaction might differ and result in varying data quality (Brand et al., 2021).

Accuracy across the screen

Accuracy across the screen in eye tracking refers to the consistency of gaze measurements across different areas or locations on the screen. It assesses whether the eye tracking system maintains accurate measurements regardless of where on the screen the participant is looking. While high-end eye tracking devices such as the Tobii Pro Fusion are reported to have consistent accuracy across the screen (Tobii Pro Fusion Data Quality Test Report, 2020), accuracy is found to be worse in the corners of the stimulus monitor (Holmqvist & Andersson, 2017). This might depend on the individual participant (Holmqvist & Andersson, 2017) but can be expected to be problematic for YET as it functions with only one camera, asymmetrically.

Light conditions

Light is essential in the recording of the eye and the estimation of its gaze. Adequate lighting enhances the visibility of the eye's features, such as the pupil. Generally, a well-lit room is recommended while direct or ambient sunlight should be avoided (Holmqvist & Andersson, 2017). A dark room affects the pupil in its size and variability; therefore, a bright room is recommended for better data quality (Holmqvist & Andersson, 2017). In this study we compared three conditions of light attached to the head mount pointed to the recorded eye to evaluate their impact on YET´s performance. A light attached to the head mount should ensure a consistent image of the eye while compensating for head movements.

The integrated LED of the endoscope camera was used in one condition. LED lighting provides a direct, visible source of illumination that enhances the contrast between the pupil and the sclera, making it easier for the camera to capture detailed eye images. However, it may cause reflections or glare. LED pointed straight to the eye should reduce shadows cast by the eyelids or eye socket. By providing a stable and even light source, LEDs help to reduce variability caused by uneven illumination, ensuring that the eye tracker can capture consistent data regardless of slight movements or changes in head position. At last, LED lighting is cost-effective and widely available making it a practical choice for a low-budget do-it-yourself eye tracker.

In the other condition, a near-infrared (NIR) light attached to the camera and pointed at the eye was tested. The advantage of NIR lies in its ability to provide a strong and stable light source that is not intrusive to the user. This stability refers to the consistent intensity of illumination, ensuring that the eye is uniformly lit without fluctuations that could affect image clarity. This consistency should reduce the impact of erratic reflections and shadows, leading to clearer eye images and improved accuracy.

The third condition involved turning off both the LED and NIR lights, relying solely on the ceiling light and light emitted from the screen. This setup aimed to evaluate YET’s performance without additional artificial lighting sources attached to the head mount.

Background conditions

In addition to the varying light conditions, we tested two background conditions either being congruent or incongruent to the initial calibration screen. The first background condition entailed a white background, congruent to the background employed during the initial calibration screen before every trial. The other background condition was a black background, incongruent to the white calibration screen. This change in background color is expected to affect the illumination of the eye, potentially impacting the quality of the eye images captured by the tracker. The incongruent black background introduces a variable that could affect the ability to accurately track eye movements, as it may alter the calibration parameters. The incongruent black background might introduce calibration issues if the eye tracker is optimized for a white background. The shift in background color could affect accurate calibration, potentially leading to deviations in gaze measurement accuracy and precision. Ensuring that the eye tracker can adapt to different background colors without significant performance degradation is important for its overall effectiveness.Light attached to the head mount should negate the effect of a reduction of screen brightness. Therefore, it can be expected that the white, congruent, background should perform better in terms of accuracy and precision than the black background whereas this effect should be lower for the IR and LED conditions.

Benchmarks for Evaluating YET’s Performance

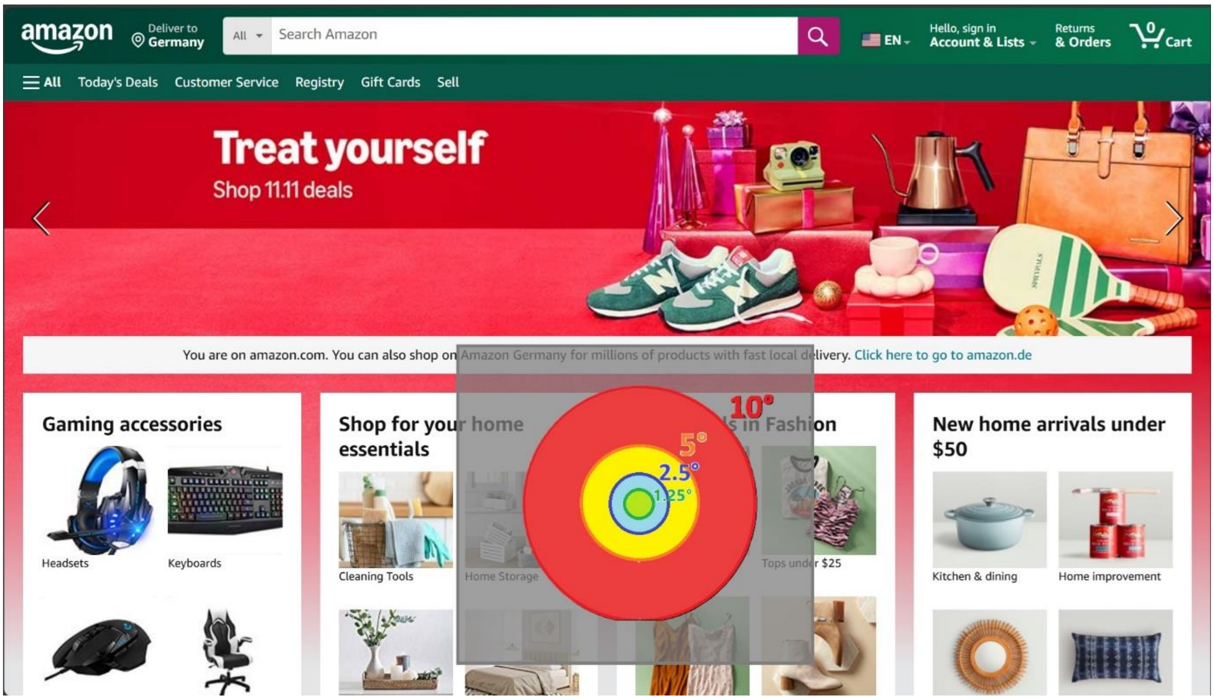
The evaluation of YET will be conducted using a series of benchmarks that compare its accuracy, precision, and consistency across various conditions. These benchmarks will serve as the standards against which YET's capabilities are assessed.

Accuracy Benchmark

The visual angle accuracy of YET will need to be evaluated to determine potential cases of application. To illustrate this, we will refer to Figure 2, which depicts a webpage displayed with circles representing various visual angles in degrees. Figure 2 shows visual angles ranging from 1.25° (green) to 10° (red) on a webpage presented on a screen with a resolution of 1920x1080 pixels, measuring 544mm x 306mm, viewed from a distance of 600mm. To effectively discern between objects on the webpage, the accuracy should not exceed the radius of the object in question. Thus, larger areas of interest such as advertisement banners, navigation bars, or main content can typically be discerned with an accuracy of 5° or below. A visual angle accuracy of 2.5° or lower allows for the identification of hyperlinks within the main content or navigation bar. Furthermore, a visual angle accuracy of 1.25° or below is required to differentiate attention among single words presented on the webpage.

Figure 2

Example of visual angles (1.25°-10°) on a webpage at 1920x1080 with screen dimensions of 544mm x 306mm at a viewing distance of 600mm



You are on amazon.com. You can also shop on Amazon Germany for millions of products with fast local delivery. Click here to go to amazon.de

However, it is important to note that the maximum accuracy needed to distinguish between objects can vary based on several factors, including the dimensions of the webpage, the resolution of the screen, and the distance from which the screen is viewed. There does not appear to be a standardized way to determine the maximum accuracy needed to differentiate e.g. between words on a webpage.

Tasks that involve fine-grained analysis of eye movements, such as distinguishing between individual words in reading research or analyzing specific fixations on small objects in cognitive tasks, demand high accuracy (≤ 1.25°).

An example could be studies that track word-by-word reading behaviour requiring high accuracy. In this context, an error margin of more than 1.25° could lead to misinterpretation of which word or letter is being fixated. Another example could be usability studies for webpages that require a differentiation between a participants focus on small objects, such as buttons or icons in a user interface. In these cases, precise identification of the exact element being looked at is essential for understanding user interaction patterns.

Moderate accuracy (1.25° to 2.5°) is typically acceptable in scenarios where the focus is on regions of interest rather than pinpointing exact details such as words. For example, research that investigates attention across broader areas such as sentences or paragraphs, images, or diagrams, may not require the same level of accuracy as studies focusing on words or buttons. An accuracy of 2.5° or below would still allow researchers to determine whether someone is focusing on a specific section of a screen or text block, enabling insights into general attention patterns. Similarly, in usability testing of digital interfaces, distinguishing between different sections, like navigation menus versus the main content area, can be achieved with moderate accuracy, as the goal is often to understand general user behaviour rather than precise fixation estimation.

Low accuracy (2.5° to 5°) is sufficient for applications that involve broader attention monitoring, where the goal is to identify general areas of visual engagement. Consumer research, for instance, often looks at which sections of a webpage attract attention, without needing exact gaze points. Similarly, different advertisements or labelling designs could be compared to understand which attracts attention first. In such cases, an error margin of up to 5° would still enable effective differentiation between large areas of interest, such as whether a viewer is focusing on an image or a block of text. Similarly, in large-scale attention studies in educational settings, where researchers aim to track whether students are looking at a general area, such as a diagram versus accompanying text on a slide, an accuracy threshold of 5° or below is adequate to capture relevant data.

If accuracy is negatively affected over the course of a trial, this could be caused by the head mount slipping out of position relative to the recorded eye (Holmqvist & Andersson, 2017). Other explanations could be fatigue or eye strain that affect the participants ability to focus.

Accuracy across the screen

While the accuracy and precision of high-end eye trackers such as the Tobii Pro Fusion is said to be consistent across the whole screen according to its data quality test report by Tobii in 2020, it can be expected that this might not be the case for YET. Unlike the Tobii Pro Fusion, which uses binocular tracking (tracking both eyes), YET employs monocular tracking with only a single, lower-quality camera. This setup may lead to greater variability in data accuracy, particularly at the screen's edges, where binocular systems typically offer improved precision by comparing inputs from both eyes. YET may perform better in terms of accuracy and precision in the center of the screen. In contrast, performance might degrade towards the edges due to the asymmetric eye recognition of solely the right eye. One side of the visual field might therefore score worse on terms of accuracy or precision.

Similar previous studies used around 20 Participants with 252 data points in total (Housholder et al., 2021; Tobii Pro Fusion Data Quality Test Report, 2020) but mentioned that more would be needed to attain more accurate results, particularly for the accuracy across the screen (Housholder et al., 2021).

Precision Benchmark

Similar to accuracy, use cases where distinguishing broad visual attention is more important than tracking fine-grained gaze estimation may allow for lower precision**.** Various combinations of accuracy and precision can significantly impact the utility and interpretation of the data.

When an eye tracking system is both accurate and precise, it consistently measures gaze positions that closely match the actual points of gaze, with minimal variability in repeated measurements. This combination represents the ideal performance allowing for reliable and valid data collection. Use cases demanding high-precision, such as reading research or detailed cognitive experiments require gaze positions to be both correct and consistently measured.

Accurate but not precise performance suggests that gaze positions are correctly identified on average but show variability in repeated measurements. This would imply that while YET can reliably approximate where the user is looking, there is noticeable inconsistency in how these positions are recorded. A potential explanation for this example could be movements of the head or the camera. If the user's head shifts slightly or if the camera experiences minor vibrations, these factors can introduce variability in the gaze measurements. Such movements may cause slight misalignments between the camera's view and the participant's eyes, leading to fluctuating recorded positions even though the accuracy remains accurate.

Inaccurate but precise performance indicate that YET measures gaze positions with minimal variability, but the measurements do not correspond accurately to the actual points of gaze. This combination implies that gaze positions are reliably measured but are systematically off from the true gaze points. For example, YET might show consistent measurements within a fixed error margin, but if the entire screen's calibration is shifted or incorrect, the data will not reflect the actual gaze positions. Due to the asymmetrical recording of one eye, this might be the case for the outer edges of the screen. Another explanation could be that a shift of the calibration takes place due to problems with the recognition of the pupil the further it moves away from the centered position. When the pupil shifts it might be partially concealed by the eyelid and at the same time more of the sclera gets exposed the further the pupil shifts from the center. Although the measurements might be consistent, they are not correct, leading to errors in data interpretation. Overall inaccurate but precise measurements could indicate that the calibration process needs refinement. If YET consistently shows high precision but inaccurate gaze positions, recalibrating the system or adjusting for environmental factors (e.g., lighting conditions, camera angle) could improve accuracy.

Both inaccurate and imprecise results would imply that YET fails to provide consistent or correct gaze measurements, showing high variability and significant deviations from the actual gaze points. If YET demonstrates both systematic errors in gaze measurement and variability in repeated trials, it will be challenging to draw reliable conclusions from the data.

Robustness

Holmqvist and Andersson (2017) define robustness as the capability of an eye tracker to perform effectively across a diverse range of participants. Variations in individual eye physiology, such as differences in eye color, can affect the quality of the data produced by the eye tracker, with some devices potentially yielding less accurate results for certain eye colors (Holmqvist & Andersson, 2017). Additionally, the depth of the eye socket or the shape and size of the eyelid might create varying shades and affect the calibration. Holmqvist and Andersson (2017) propose that an eye-tracker with the ability of varying the angle of the camera to the eye can resolve these problems and enhance robustness.

Additionally, YET makes use of a simplistic approach to predict gaze positions which may limit performance metrics compared to more sophisticated gaze estimation algorithms. Calibration tests should highlight potential flaws and identify the specific conditions under which YET performs optimally.

Methods

Study Design

The experiment was conducted in the Serious Gaming Room of the BMS Lab at the University of Twente in August 2024. The rooms of the laboratory provided conditions for the study to be conducted in a controlled and undisturbed environment. The room was equipped with a height adjustable desk and chair. A Legion Pro 5 laptop with a 16inch 240 Hz screen was used at full brightness (496 nits) to display the experiment.

YET was tested as a do-it-yourself setup with minimal requirements and barriers to potential users. A large paper towel roll was used as a headrest and the head mount for the eye tracker consisted of a cheap pair of headphones, a 30cm plastic ruler, a small 3D printed part (Clip-Y) and an endoscope camera which was connected to the laptop (Figure 3).

The Clip-Y was glued to the ruler to hold the camera. Hook and loop fastener were used on the side of the pair of headphones and the ruler to be able to attach the ruler and position the camera angle for each participant, as suggested by Holmqvist and Andersson (2017). A near-infrared light was glued on top of the eye tracker and connected to the laptop.

To create consistent lighting conditions, the blinds of the rooms were shut while the lights were turned on at maximum brightness. To minimize shadows cast from the ceiling lights of the room on the face of the participant, the desk was moved to the middle of the room where ceiling lights were situated more in front of the participant. The headphones and the desk were disinfected in between participants and a paper towel was placed on the head rest to ensure hygiene.

Figure 3

YET Apparatus and experiment setup

Figure 4

Example of YET apparatus in use

The viewing distance of the screen to the participant was set at 60cm. Each participant took part in 3 trials per lighting condition. One trial consisted of 48 stimulus positions (24 for each background) that were randomized. The lighting condition differed based on whether the near-infrared light was turned on (Condition 1), or whether the integrated LED of the endoscope camera was used (Condition 2), or not (Condition 3). Each potential order of conditions (6 in total) was tested four times. In total, each participant took part in 9 trials (3 light conditions x 3 trials) with each trial consisting of 48 stimuli (24 calibration points x 2 background conditions). Per condition, each participant was tested on 72 data points (3 trials x 24 calibration points x 1 light condition x 1 background condition). In total, 10368 data points (24 participants x 3 trials x 24 calibration points x 3 light conditions x 2 background conditions) were tested.

Procedure

Once a participant arrived, the purpose and procedure of the study was explained orally. Participants were informed about the risk and advised to take breaks in between trials and to look out of the window every 20 minutes to minimize eye strain. Further, participants were reminded that they could withdraw at any time. Then they were handed an information sheet and informed consent form (Appendix A).

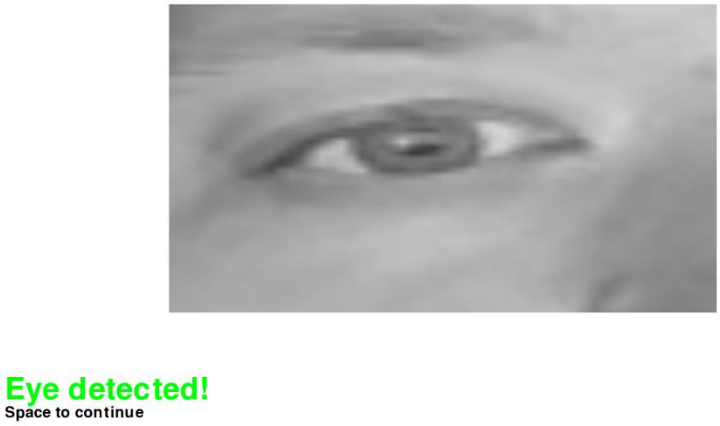
After signing the informed consent, participants were seated at the desk with their chin positioned on the head rest. Then they were asked to adjust the height of the chair and the height of the table to minimize the need to lean forward. Once an ergonomic position was found, the laptop was moved 60cm away from the participant.

Each participant was informed about the 9-point calibration at the start of every trial. After, they were asked to keep as still as possible throughout each trial. It was emphasized that the eyes should remain in a neutral position (not widely opened) and that the head should not move, nor should it be tilted. Further, participants were asked to close their jaw with their teeth touching each other while resting on the headrest.

The headphone was then put on as tight as possible, and the participant was asked to find a position that could be sustained throughout the experiment. Once ready, the software was started for the first trial, beginning with the recognition of the eye (Figure 5). This part merely assists the user of the eye tracker in assuring that the eye is recognized and thereby prevents utter misuse. If the eye could not be recognized immediately, the ruler was detached and adjusted for the camera to be able to attain a better image of the eye, before repositioning.

Figure 5

Eye recognition



Eye detected!
Space to continue

The participant was then instructed to press space. At the first trial this was used to prime the participants to maintain their position while pressing space later during the experiment. The 9-point calibration test was then started (Figure 6). After the calibration, a screen appeared with a red circle at the position where the software estimated the participant´s gaze (Figure 7). The researcher then pointed to the middle of the screen and asked the participant to follow the finger from the middle to each direction. If the circle was far off (4°) another calibration test was tried.

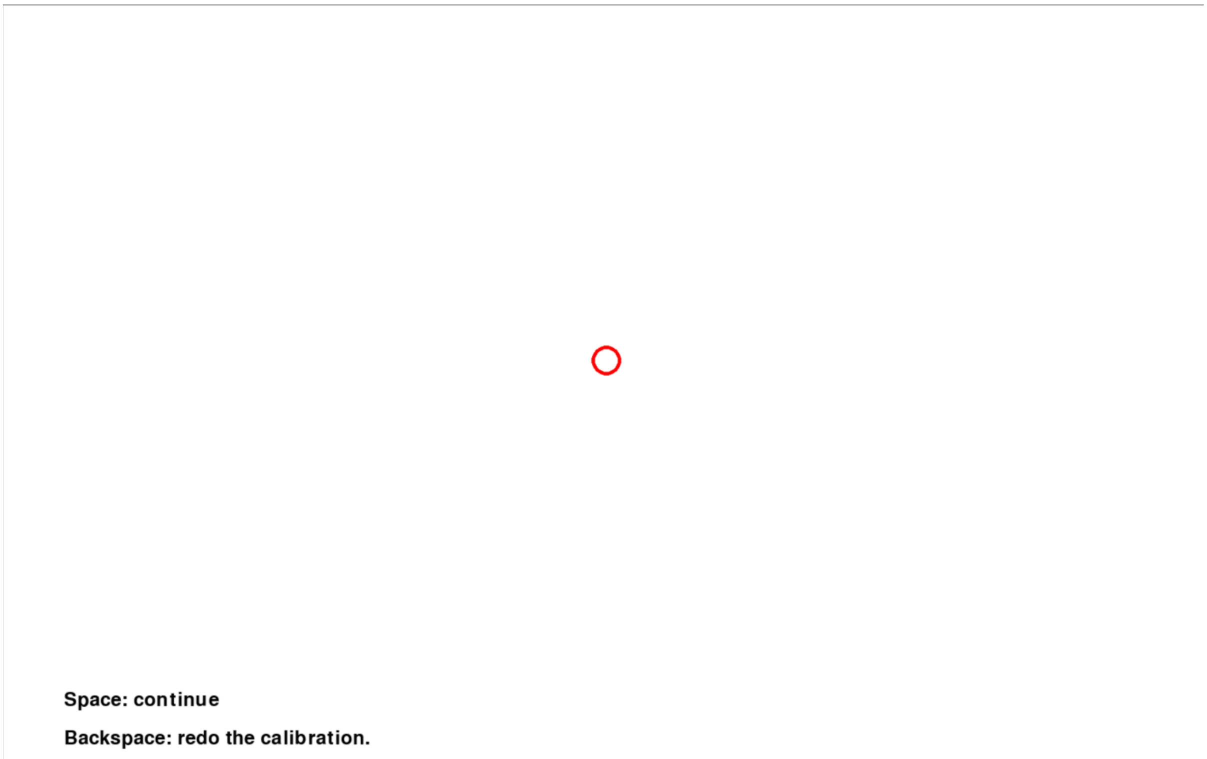
Figure 6

Calibration Screen

Focus on the orange circle and press Space.

Figure 7

Verification of calibration



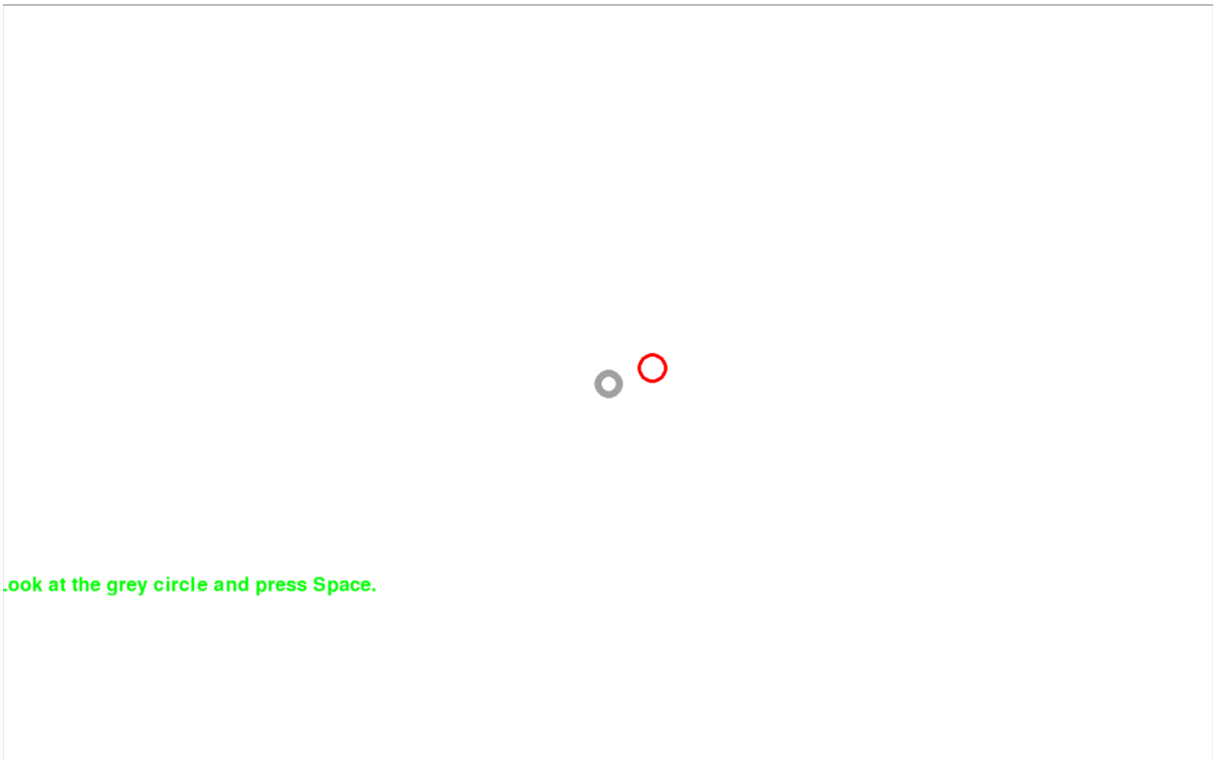
Space: continue

Backspace: redo the calibration.

Once the calibration was successful, participants were instructed to look at similar calibration stimuli spatially arranged in a 24-point grid (4x6). Each of the 24 stimuli was presented with a white and black background in a randomized order to examine the effect of the screen lighting on measures. Before each stimulus, participants were instructed to look at a circle at the center of the screen and proceed by pressing space (Figure 8).

Figure 8

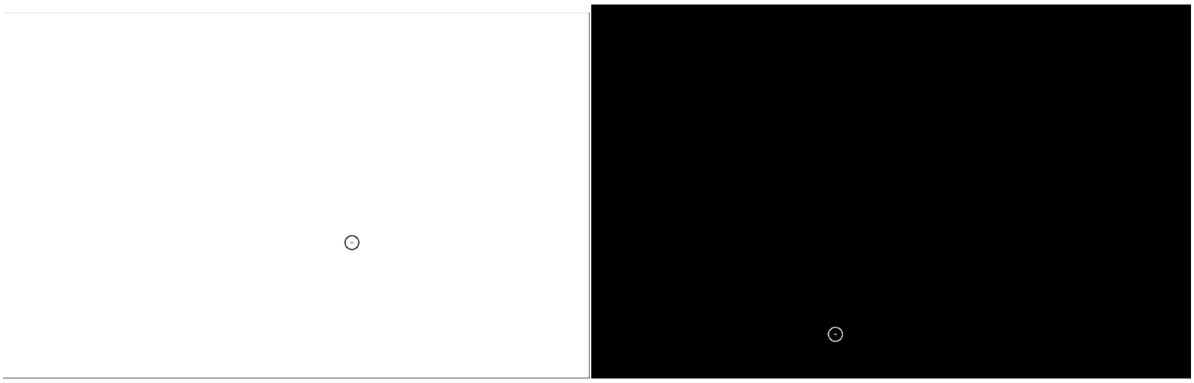
Recalibration Point



ook at the grey circle and press Space.

Figure 9

Examples of Stimulus with white and black background



Participants were given the option to take a break to rest their eyes after completion of a trial. After every 20 minutes, participants were again reminded and given the option to look out of the window to reduce eye strain.

Participants

24 participants took part in the study. The participants were recruited via different forms of media, such as social media, communication applications and flyers. Participants with glasses, lenses or any type of sight corrections were excluded from the study.

Data analysis

Data analysis was conducted under two varying conditions, including two different background colors (white and black) and a different light source attached to the head mount (NIR vs LED vs noLight).

Accuracy refers to the closeness of estimated participant responses to the true target locations and was measured by comparing the estimated gaze coordinates with the coordinates of the intended target stimulus. Accuracy was calculated as the average deviation between the actual and intended gaze points. Specifically, the Euclidean distance between the actual gaze point and the center of the target stimulus was computed. The distance in pixels from each estimated gaze point to the target stimulus was converted to angles based on the screen parameters and viewing distance before calculating means. Lower values indicate higher accuracy.

Precision refers to the consistency or repeatability of participant responses and was measured by assessing the variability in gaze points across repeated trials. Precision was calculated as the standard deviation of the gaze points around the mean gaze point for each participant. Lower standard deviation values indicated higher precision.

Both accuracy and precision were calculated at a participant level and group level for each of the conditions. Gaze plots were then created by inverting values and mapping the estimated gaze points and stimulus positions. These gaze plots were then examined and compared for each participant per trial per background condition. At the center of each gaze plot, gaze points were found. To further understand these, graphs were created to show the increase in accuracy based on the horizontal and vertical shift.

By excluding the first 10 observations for every stimulus per trial per participant, the points at the center of the gaze plots and a bulge in the middle of the graphs showing the increase in accuracy based on the horizontal and vertical shift was reduced. Gaze plots and accuracy and precision were examined again after exclusion of the first 10 observations for every stimulus per trial per participant.

At last, the accuracy across the screen, per stimulus, was examined. For this, plots were created showing circles representing the mean Euclidian distance around each stimulus.

To make an estimation of the accuracy given a certain field of view, stimuli were excluded. The first and last column of four stimuli was excluded first, then the upper and lower row as well and at last only the center four stimuli were observed.

Generalized Linear Mixed Model

To evaluate the accuracy of the eye-tracker system under different lighting and background conditions, we employed a Generalized Linear Mixed Model (GLMM) with both fixed and random effects. This model allows for the analysis of overall trends while accounting for individual variability in participants and target locations.

Accuracy ~ LightCondition * background + (LightCondition * background|Participant) + (LightCondition * background|Stim)

Accuracy was used as the dependent variable for the model. The fixed effects included two main factors: Light Condition and Background, as well as their interaction. The lighting conditions were “NIR”, “LED”, and “noLight”, and the background conditions included a white background, congruent to the background of the initial 9 grid calibration before every trial and a black background, which was incongruent.

(LightCondition * background|Participant) is the first random effects term for participants. We expected that the effect of light and background (and their interaction) might vary across participants.

(LightCondition * background|Stim) represents the second random term for stimulus target positions. Each calibration target may have different accuracy results under the various light and background conditions. For example, accuracy might be more affected by lighting in some areas of the screen than others. The model is designed to substantiate both the effect of light conditions and background on accuracy and the random factors participant and stimuli.

Results

Background condition (white vs black background)

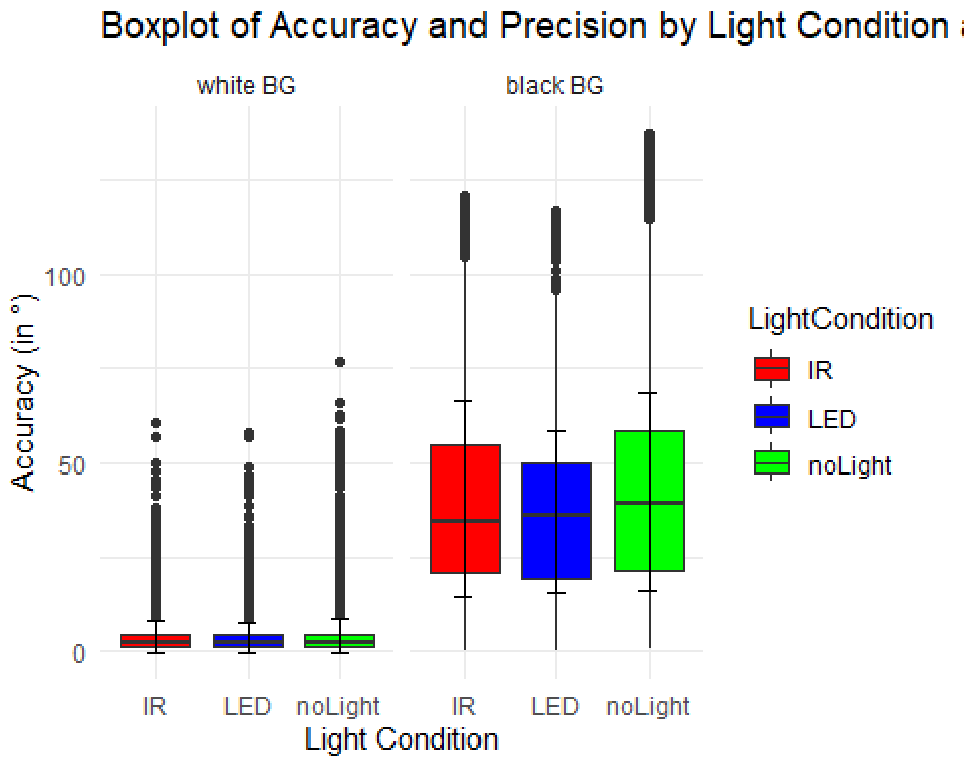
The accuracy and precision across the whole screen for the black background were not in the range of useable metrics for eye tracking with an accuracy of around 40° (see Table 1).

The boxplot below visualizes the difference in accuracy between light conditions and background. For the white background condition, the LED condition had the lowest Accuracy (3.71°), followed by NIR

(3.88°), and highest with no light (3.96°). In contrast, for the black background condition accuracy worsened beyond usable levels (NIR = 40.59°, LED = 37.01°, noLight = 42.52°).
The model predicted that changing the background from white to black results in a substantial increase in accuracy by 39.59°, with a probability of 95% that the accuracy is between 35.08° and 45.6°.

Figure 10

Boxplot of Accuracy and Precision by LightCondition



The gaze plots clearly indicate that estimated gaze points were scattered around the stimulus positions for the white background (Figure 11). For the black background, estimated gaze points did not match stimulus positions. For participants who had an overall good accuracy for the white background, it appears that the estimated gaze points were still aligned in a grid but shifted to another position (Figure 12). For participants who had a worse accuracy for the white background, the estimated gaze points were more scattered and not aligned like the stimulus grid (Figure 13). This observable shift occurred in various directions, often diagonally but sometimes straight up, down or to the left or right of the stimulus grid.

Given that the estimated gaze points for the black background condition were not scattered around the stimulus grid, their accuracy and precision values were not useable.

Figure 11

Estimated gaze points (red) and stimulus positions (black) for the first trial of IR-condition, for the white background for participant 6

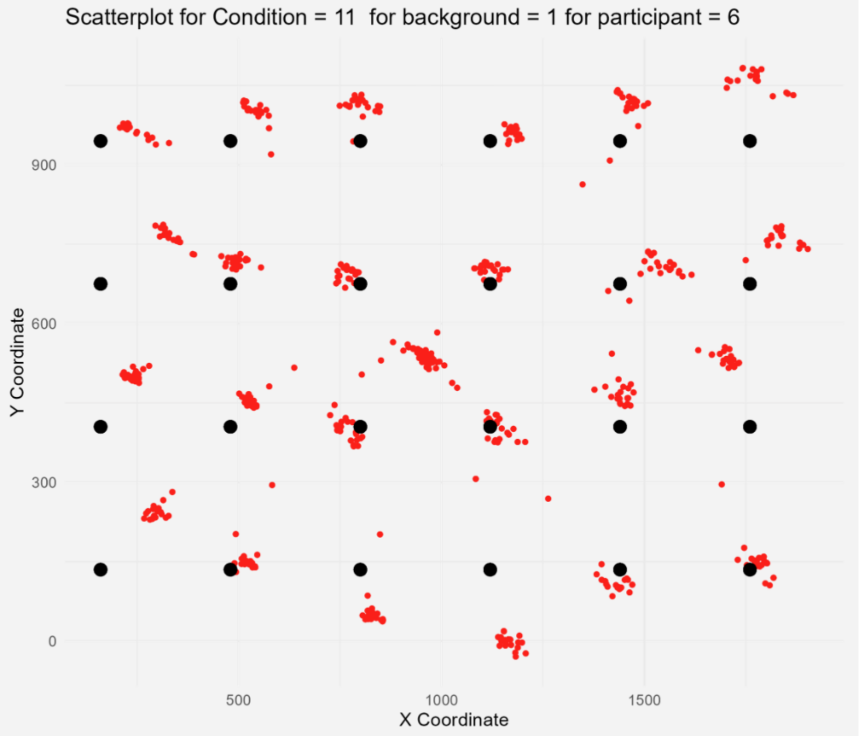


Figure 12

Estimated gaze points (red) and stimulus positions (black) for the first trial of IR-condition, for the black background for participant 6)

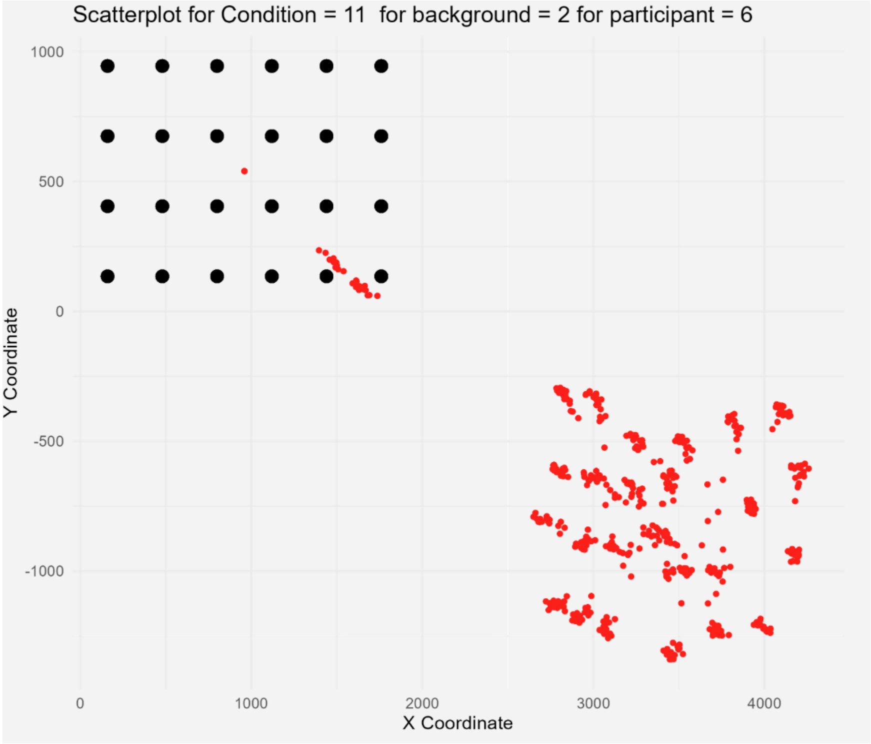
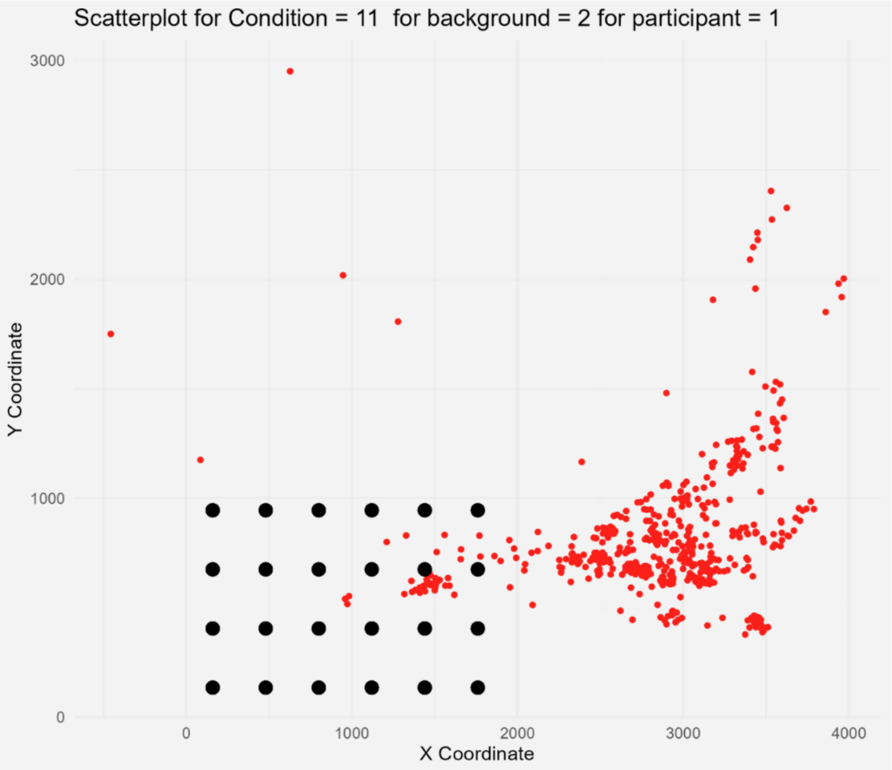


Figure 13

(Estimated gaze points (red) and stimulus positions (black) for the first trial of IR-condition, for the black background for participant 1)



Lighting conditions attached to head mount

The endoscope camera has an integrated LED light that was tested against an optional near-infrared light attached to the camera pointing to the eye. For this, three conditions were tested either with the near-infrared light turned on (“NIR”), with the LED turned on (“LED”) or with no light (“noLight”).

No noticeable differences could be found between these conditions. The mean accuracy of these conditions for the white background was similar; the accuracy for the white background across all stimuli and trials: NIR = 3.88°, LED = 3.71°, noLight = 3.96° (Table 1).
No noticeable difference in the shift of the estimated gaze points for the black background could be found between the lighting conditions.

The intercept representing the baseline accuracy under the reference condition (NIR light, white background) was estimated to be 3.07° (95% CI: [2.19°, 3.86°]). When switching to the LED light condition, a slight decrease in accuracy of 0.18° was observed (95% CI: [-0.43°, 0.07°]), though this change was not statistically significant because the credibility intervals include both positive and negative values, which means that the true effect could be an increase, a decrease, or no change at all. Similarly, in the "no light" condition, there was a negligible increase in accuracy of 0.08° (95% CI: [-0.29°, 0.51°]) with the confidence interval including zero, indicating minimal change in accuracy compared to the baseline.

Thus, across lighting conditions (NIR, LED, and noLight), as no notable differences in accuracy were found for the white background the observed differences could be due to random variation rather than effects of the conditions. Similarly, no significant differences in the shift of the estimated gaze points were observed for the black background, where accuracy was already outside usable levels regardless of lighting.

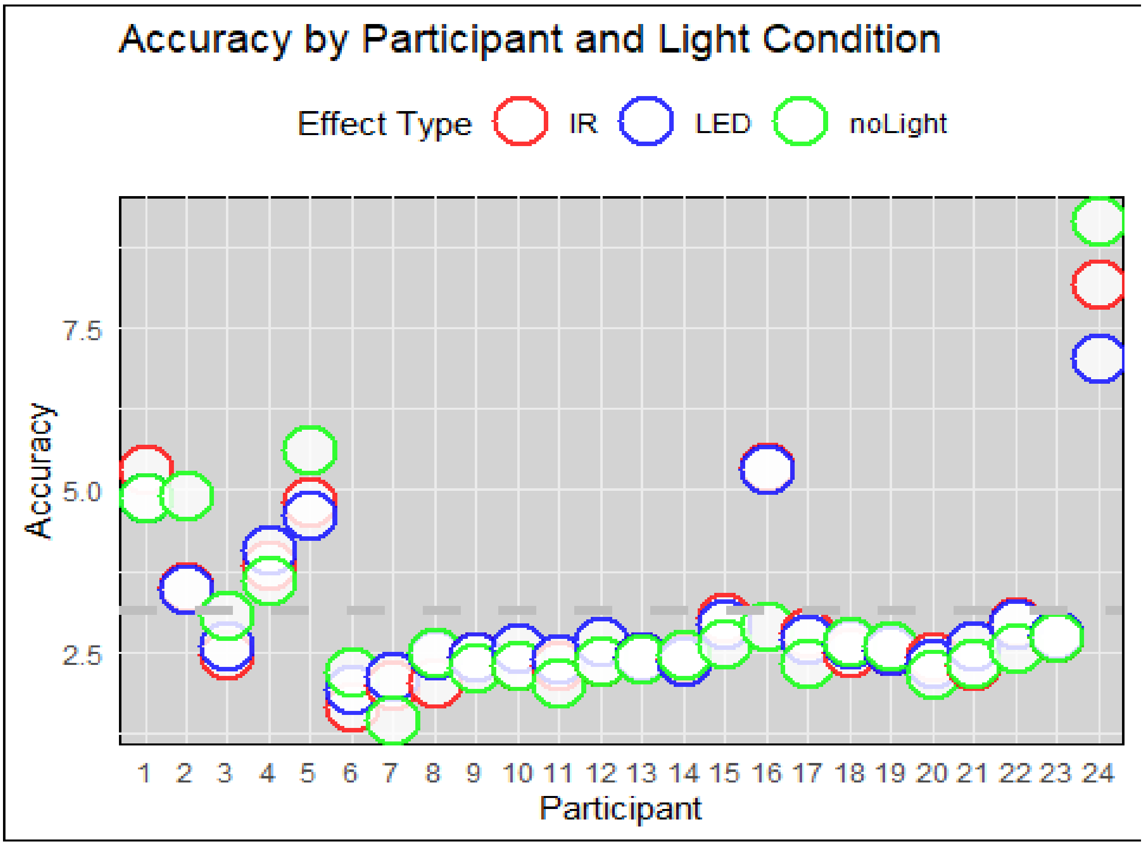
Participant-level analysis

The GLMM indicates that only 5 participants from 24 show a positive intercept (e.g., Participant 1, 4, 5, 16, 24) while 19 participants show negative values (e.g., Participant 3, 6, 7, 8, 9, 17, and 20). For example, participant 5 was predicted to show an increase of 1.68° compared to the baseline for the white background and NIR light condition. At a certainty of 95% it was predicted to be around 0.93° to 2.4°. Participant 24 had the highest increase compared to the baseline with an intercept of 5.07° and a statistically significant predictable certainty of 95% being in between 4.18° to 5.89°. Others performed better, e.g. participant 6 showed a reduction of 1.48° for which we can be certain that it lies between -2,17° and -0.71° at a confidence interval of 95%.

For all participants except for one, the accuracy was consistently similar across lighting conditions for the white background. Only participant 16 performed noticeably better on condition 3 (noLight) than compared to the other two conditions (Table 2, Figure 14). Participants generally appeared to produce similar results across light conditions (Figure 14).

Figure 14

Accuracy (in °) for Participants by Light Condition for the white background; lower values represent less visual angle error in degrees



Filtering of data

Each scatterplot for the estimated gaze positions for the white background had estimated gaze points at the center of the screen (Figure 11). This was further highlighted by creating plots that showed the increase in accuracy based on a vertical or horizontal shift (Figure 15) where a bulge could be seen at 0°. After filtering out the first 10 observations for every observed trial of each stimulus position per participant, per condition, per background, the estimated gaze points for the scatterplots (Figure 17) and the bulge in the plots for the increase of accuracy per vertical or horizontal shift (Figure 16), reduced.

The accuracy and precision changed noticeably across all conditions after filtering out the first 10 observations; IR = 3.09°, LED = 2.9°, noLight = 3.17° (see Table 3).

Figure 15

Accuracy for horizontal and vertical shift, including 95% of the data.

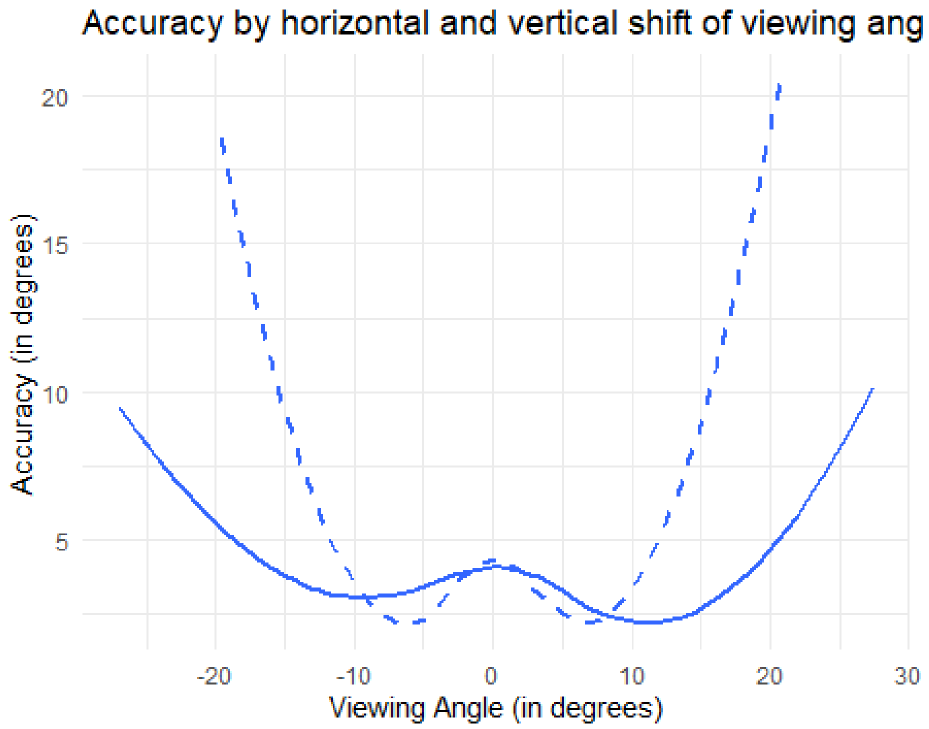


Figure 16

Accuracy for horizontal and vertical shift, including 95% of the data. Excluding first 10 observations; the dashed line represents the vertical shift and the solid line the horizontal shift

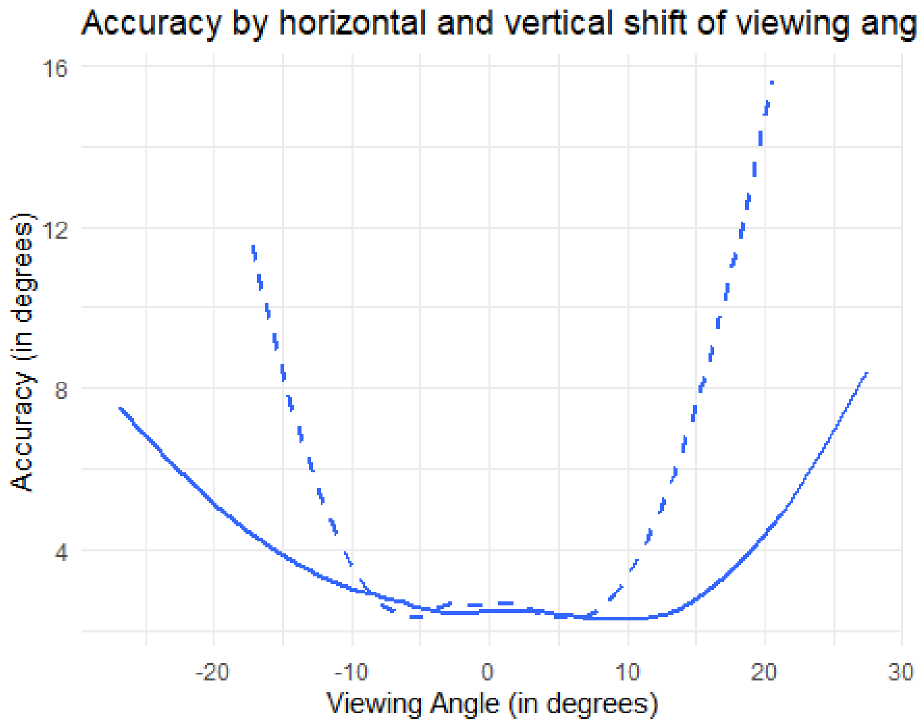
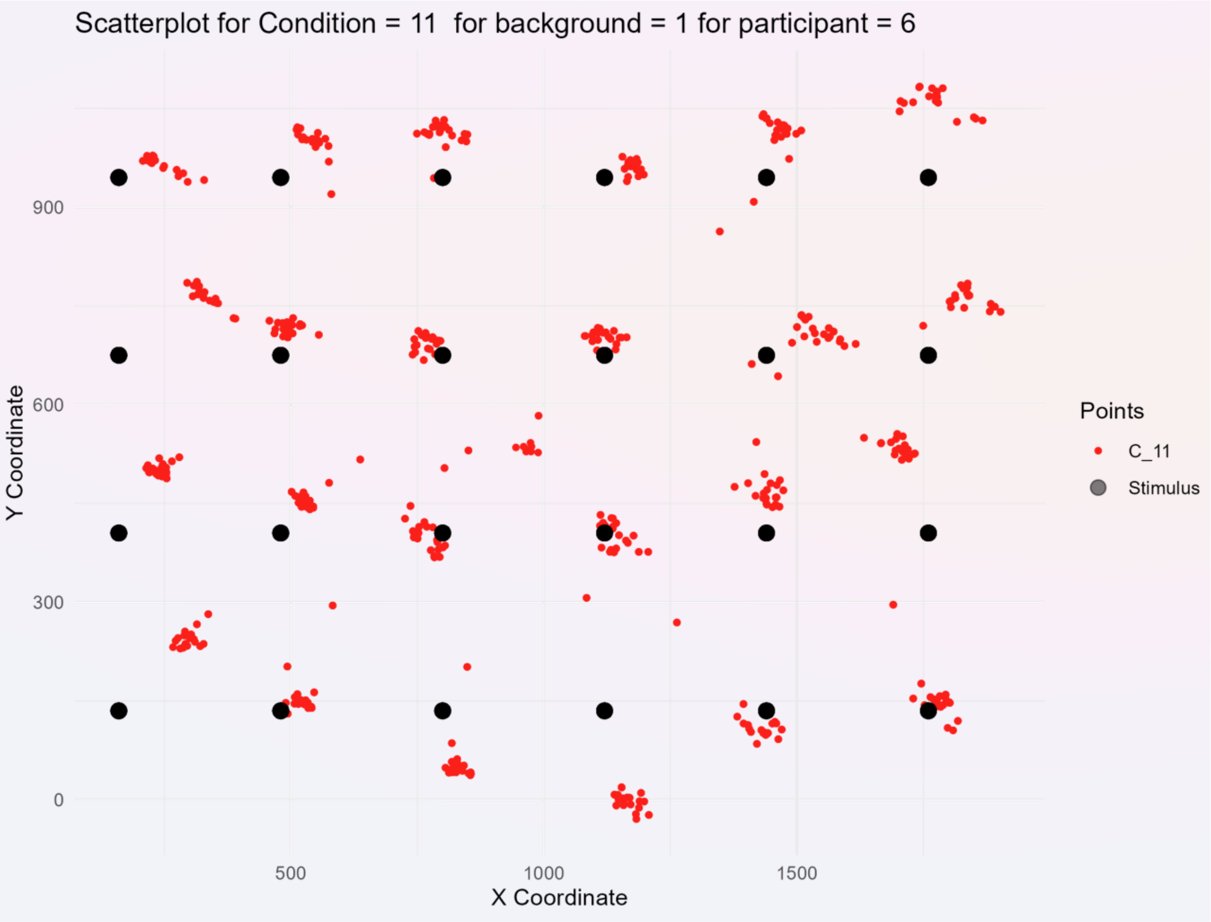


Figure 17

Scatterplot for first trial of NIR-Condition for the white background for P6 after filtering the first 10 observations; the dashed line represents the vertical shift and the solid line the horizontal shift



Accuracy across the screen

The accuracy was computed for each stimulus and a plot was created to highlight the different mean Euclidian distances per stimulus per lighting condition for the white background (Figure 18). The further away from the center of the screen, the worse the accuracy of each stimulus position was (Figure 18). After excluding the left and right row of stimuli, the accuracy changed to: IR = 2.55°, LED = 2.46°, noLight = 2.5° (see Table 4). When only the four stimuli at the center of the screen were included, the accuracy improved further to: IR = 2.15°, LED = 2.14°, noLight = 2.17°. The stimuli located at 3.69° around the center (10, 11, 14, 15) had the best accuracy, represented by negative intercept values. Additionally, the outer left row appears to show larger circles than the last row on the right.

Figure 18

Stimulus Positions with Mean Euclidian Distance Circles by Condition Group, for the white background.

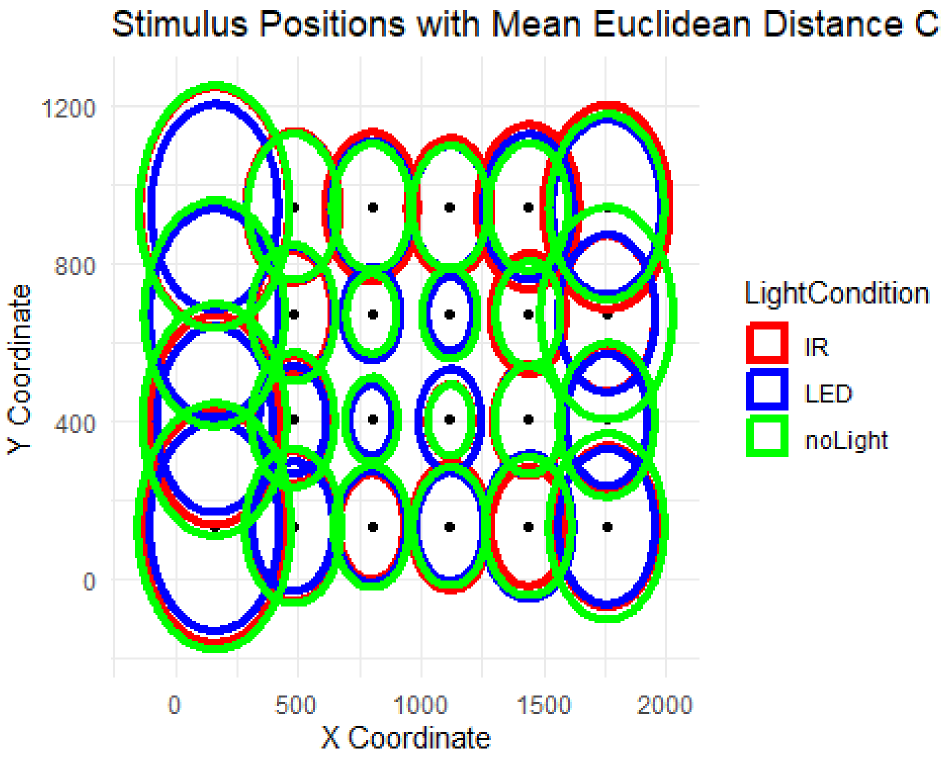
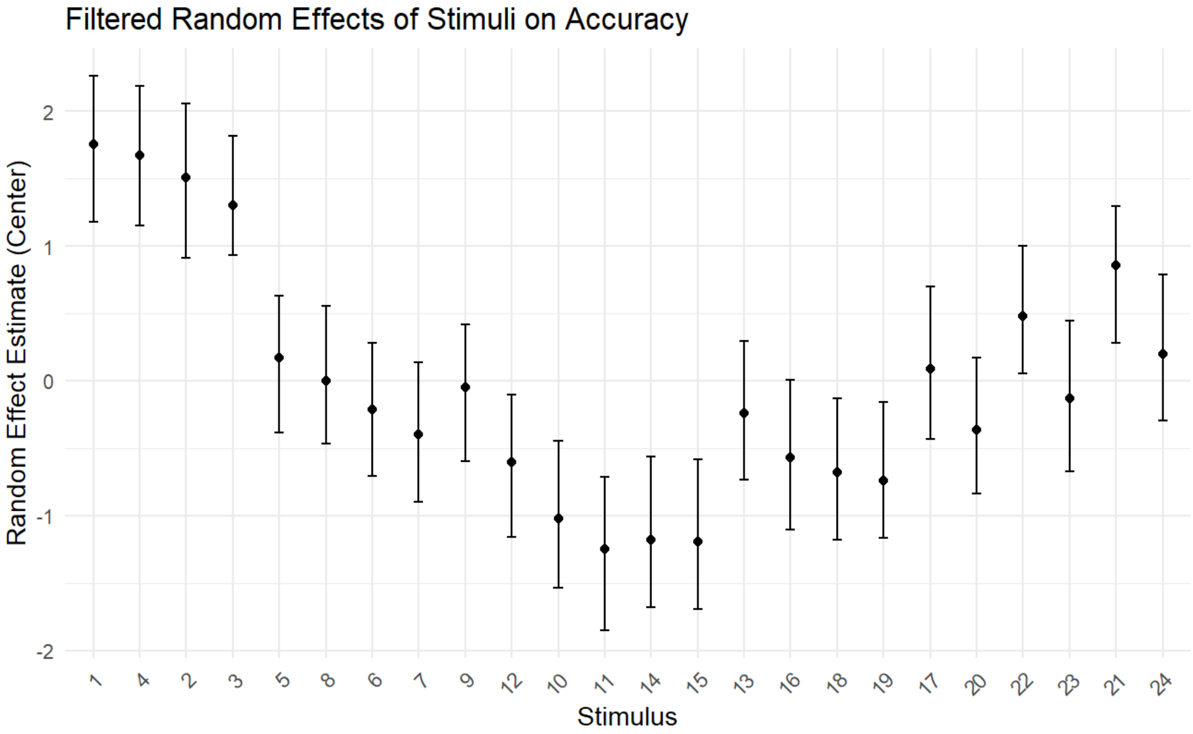


Figure 19

Random Effects of Stimuli on Accuracy for the white background; the 24 Stimuli are ordered by their distance in visual degrees to the center.



Ordering of stimuli for Figure 19 according to the 4x6 grid:

1	5	9	13	17	21
2	6	10	14	18	22
3	7	11	15	19	23
4	8	12	16	20	24

Stimuli 10, 11, 14, and 15 were closest to the center with a distance of 3.69°. For example, for Stimulus 10, the intercept value is -1.02, and we can be 95% confident that the true intercept is between -1.53 and -0.44. All four stimuli located at around the center of the screen showed statistically significant decrease in accuracy, suggesting that there is strong evidence that the accuracy at these stimulus positions is not simply due to chance.

The stimuli located 7.66° from the center of the screen (Stimuli 9, 12, 13, 16) all showed a decrease in accuracy albeit not being statistically significant, except for stimulus 12 with a reduction in accuracy of -0.60° compared to the baseline. The confidence interval for Stimulus 12 ranged from -1.16° to -0.10°, indicating that this decrease in accuracy is statistically significant at the 95% confidence level.

The other stimuli at this distance (Stimuli 9, 13, and 16) did not show statistically significant differences, as their confidence intervals included zero, with stimulus 16 (Intercept = -0.57, 95% CI [-1.11, 0.01]) almost excluding zero.

At 8.77° from the center of the screen stimuli (6, 7, 18, and 19) all displayed a decrease in accuracy, while only the stimuli on the right side, 18 and 19, showing statistical significance. Stimulus 18 had an intercept of -0.68, with a confidence interval of -1.18° to -0.13°. The fact that this confidence interval does not contain zero indicates a statistically significant reduction in accuracy. Similarly, stimulus 19 was predicted to have a reduction of -0.74 from the baseline, with a confidence interval of -1.17° to -0.16°.

None of the stimuli (5, 8, 17, and 20), located 11.03° to the screen´s center, showed statistically significant reductions in accuracy. Stimulus 8 and 20 were predicted to have a slight reduction of accuracy (Stimulus 8: Intercept = -0.001, Stimulus 20: Intercept = -0.37) while stimulus 8 is expected to lie at the interval of -0.47° to 0.55° with a certainty of 95% and stimulus 20 between -0.83° to 0.17°. In summary, all stimuli located 11.03° from the center of the screen demonstrated non-significant reductions in accuracy. This suggests that at this distance from the center, there is insufficient evidence to conclude that accuracy is significantly different from the baseline.

The stimuli located at the border of the screen (Stimuli 2, 3, 22, and 23), each positioned 14.22° from the center, exhibited notable variations in accuracy, with some demonstrating significant positive increases in accuracy. Stimulus 2 had an intercept of 1.51, with a confidence interval ranging from 0.91° to 2.06°. This confidence interval does not include zero, indicating that the intercept for Stimulus 2 is statistically significant. This suggests that accuracy at this stimulus position is significantly higher than the baseline. Similarly, stimulus 3 (Intercept = 1.31, 95% CI [0.93, 1.82]) and stimulus 22 (Intercept = 0.48, 95% CI [0.05, 1.0]) both indicated statistically significant effects. Stimulus 23 recorded an intercept of -0.13, with a confidence interval of -0.67° to 0.44°. The confidence interval for Stimulus 23 includes zero, indicating that this intercept is not statistically significant. This result suggests that the accuracy for Stimulus 23 does not differ significantly from the baseline.

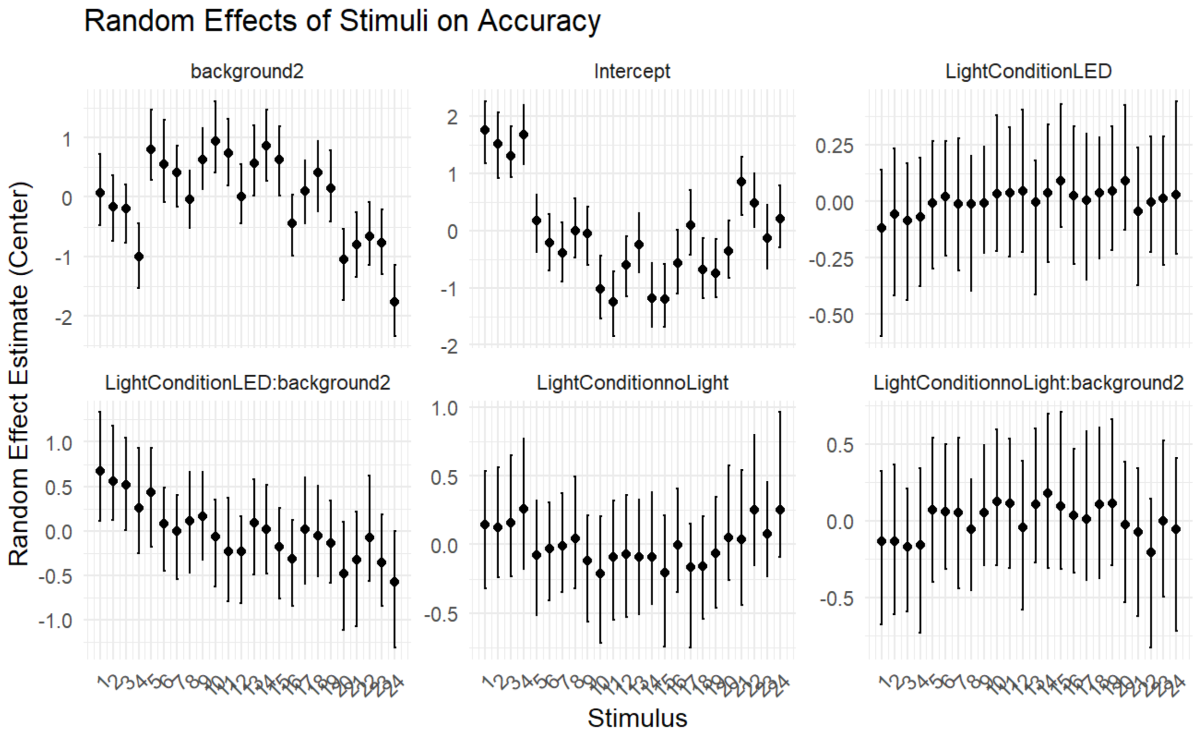
Stimuli (1, 4, 21, and 24) are located at the corners of the screen, and thereby furthest away from the center, with 15.7°. Stimulus 1 at the upper left corner of the screen and stimulus 24 at the lower right corner of the screen both showed statistically significant increases in accuracy compared to the baseline. Stimulus 1 had an intercept of 2.29, with a confidence interval ranging from 1.46° to 3.24°. Stimulus 24 recorded an intercept of 5.15, with a confidence interval of 4.28° to 6.08°. Stimuli 4 and 21 did not demonstrate significant differences from the baseline.

As expected from the previous results on accuracy, the incongruent, black, background condition consistently exhibits a positive effect on accuracy across most stimuli. For example, for Stimulus 10, the effect of background2 resulted in an intercept of 0.9330 (CI: 0.38 to 1.58), suggesting that this stimulus under the background2 condition significantly increases accuracy. However, some stimuli, like Stimulus 24, showed a reduction (-1.7734, CI: -2.40 to -1.21).

The analysis of the Light Conditions (LED and noLight) presents a more mixed picture; for Stimulus 1, the LED condition exhibited a negative effect with an intercept of -0.0992 (CI: -0.54 to 0.16). The noLight condition showed a non-significant positive effect with an intercept of 0.1199 (CI: -0.32 to 0.56). Stimulus 11 reflected a similarly non-significant trend, with the LED condition at 0.0335 (CI: -0.26 to 0.39) and noLight at -0.1225 (CI: -0.48 to 0.31). In summary, the Light Conditions do not exhibit a strong, consistent pattern of effects on accuracy across the stimuli.

Figure 20

Random Effects of Stimuli on Accuracy



Discussion

The present study aimed to evaluate the feasibility of YET (Your Eye Tracker) as an ultra-low-budget, do-it-yourself eye-tracking solution. By systematically comparing the accuracy and precision of YET under varying conditions, we sought to understand its potential as a viable alternative to more expensive, commercially available eye-tracking systems. We employed a within-subjects calibration study with three varying light conditions attached to the head mount of the eye tracker, namely an infrared light (IR), an LED light and no light for the third condition. Furthermore, we tested two varying background conditions in each trial consisting of a white background congruent to the background during the initial calibration at the start of every trial, and an incongruent black background.

YET demonstrated feasibility as a low-cost do-it-yourself eye-tracking solution, with limitations under certain conditions. For trials conducted with the white background, which was consistent with the calibration background, the system showed moderate accuracy and precision across the screen. Specifically, YET achieved an overall accuracy of approximately 3.85° across the entire screen, with precise estimations (around 2.17°) at the center of the screen. This suggests that under ideal conditions—such as matching background brightness and central gaze targets—YET is capable of achieving a comparable level of accuracy to other established and cost extensive eye tracking systems with a reported accuracy of 2° and lower.

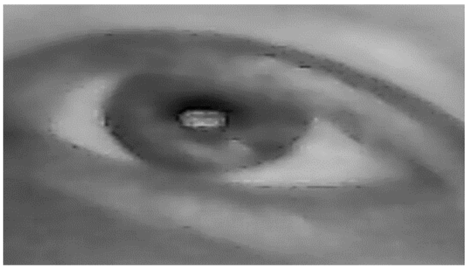
Discussion of Results

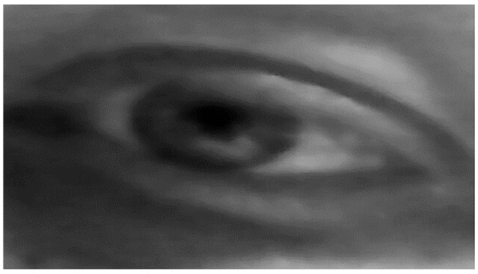
White background vs black background

The gaze points matched and were scattered around the stimulus positions for the white background (Figure 11). Gaze points did not match stimulus positions when the background was black (Figure 12 & 13). There appears to be a shift of the calibration grid in the gaze point estimation when the background was black (Figure 12), this appeared to be the case for participants that had a good accuracy. The participants whose gaze points could not be estimated more accurately for stimulus with a white background, showed a skewed/morphed grid with a shift (Figure 13). The grid shifts diagonally or in one of four directions (up, down, left, right) most of the time, at times it is skewed. For those participants whose gaze were estimated more accurately in white backgrounds, the gaze points were more closely scattered around these shifted stimuli positions. For each trial, therefore for each calibration, the direction in which the grid changed was different. The shift was detectable across the conditions of the light attached to the head mount (“NIR”, “LED”, “noLight”). No differences in the shift could be detected between these conditions.

What happens when the background turns black?

When the background turns black, the illumination of the eye by the screen changes and the reflection of the screen in the pupil disappears. Additionally, the pupil might increase in size and shadow emerge that were hidden by illuminating the eye from the front.





There appears to be a misalignment in how the software interprets these gaze positions. The calibration was done with a white background; the software calibrated gaze positions given a different brightness condition for each quadrant when the background changes to black.

This might imply that if the brightness settings differ during the calibration to the experiment, that the gaze point estimates are not accurate and skewed into a direction.

Accuracy and lighting conditions

The overall accuracy is promising, given the early stage and potential aspects of improvements. Our analysis of different lighting conditions (near-infrared light, LED, and no light) reveals no significant differences in accuracy or precision across these conditions when using a white background. The overall accuracy (for the white, congruent background) between the conditions (“NIR”, “LED”, “noLight”) did not differ noticeably; most participants were found to have similar accuracy across the light conditions (Figure 14).

This might suggest that YET's performance is relatively stable across different lighting setups, provided there is sufficient ambient light to enable clear contrast detection. Though, further testing is required to infer potential effects of the angle of the light attached to the head mount relative to the recorded eye. Further interaction effects might occur based on the other light sources used, based on various eye shapes or eye colors. Another probable explanation could be that the light sources attached to the head mount were not ideally adjusted. Especially the infrared source can be sensitive to the angle of illumination. The LED light might eliminate some shadow especially from the ceiling lights but might create new shadows if it is positioned to the eye from one side, here more LEDs might create more optimal lighting conditions potentially even negating the effect of the incongruency of the background. This highlights the future potential of YET as a DIY eye tracker with the possibility for optimization.

In the table, the overall accuracy in degree is given with the accuracy for the filtered dataset (in brackets) where the first 10 observations of every trial where excluded. The first 10 observations were gaze points at the center of the screen that were observed before the participant reacted. Excluding these observations where participants were still looking at the center of the screen (as instructed before each stimulus position was shown) results in a more accurate estimation of the accuracy while a participant is focusing on a stimulus position.

The outer positions should be affected most by observations that were recorded before the participant reacted. Though, plot with the mean Euclidian distance per stimulus position (Figure 18) illustrate that the accuracy is still worse the further away a stimulus position is from the center of the screen.

At the center of the screen, the four calibration points at a distance of 3.69° to the center of the screen where found to result in the most accurate gaze estimations. At the corners of the screen, with 15.7° from the center of the screen and outer edges with 14.22°, the accuracy was worst (Figure 18).

In the recorded image of the eye values for the brightness of each quadrant are affected by increased changes the further away the gaze is from the center of the screen. The opening of the eyelids and change of the position of the pupil result in more of the sclera being exposed. Additionally, parts of the pupil might be hidden under the eyelid, especially when it moves up. Similarly, more of the pupil is being exposed when it moves down and was hidden below the upper eyelid. Movement of the eyelids might affect shadows of the eyelids as well. In addition, the further away from the center, the more susceptible participants might be to potential head movements following their gaze.

The accuracy is likely worse for stimulus positions that were further away from the center of the screen since these additional effects took place compared to positions at around the center of the screen.

Participant Variability

The results highlight inter-participant variability in accuracy. While most participants achieved acceptable levels of accuracy and precision, others (P1, P5, P24) performed poorly across all conditions. This variability could stem from several factors, including differences in individual eye physiology, head movement stability, and familiarity with (or adherence to) the calibration process. For instance, participant 16 improved performance under the no-light condition, as her last tested condition, could suggest that for some users, minimizing distractions or reflections may enhance the system's effectiveness. This finding points to the importance of personalized calibration procedures and suggests that YET might benefit from additional user training or adaptive algorithms to account for individual differences.

Feasability of YET

The results of the accuracy are promising with an accuracy of ~3° across the screen (after filtering for the initial first 10 observations) comprising a field of view of around 31.4°. At around ~7.4° parting from the center of the screen, accuracy was best with ~2.15° (sd = 2.8°). The accuracy indicates that at the center of the screen YET can achieve comparable accuracy to other eye tracking systems who are reported to be <2° while filtering for the initial first 10 observations but without excluding outliers. Though, YET´s precision did not match other eye tracking systems yet. In this study YET was tested under consistent light conditions but with minimal affordances to its setup. A true do-it-yourself version was tested that could be employed by anyone with a paper towel roll as a headrest and a simple head mount construction consisting of a pair of headphones costing 8€, a ruler, hook and loop fastener, a small 3D printed part and an endoscope camera costing around 10€. This setup can likely be improved with a more stable head mount consisting of a counterweight and a more secure attachment than hook and loop fastener, additional light attached to the head mount or positioned around the screen, illuminating the eyes.

One of the challenges observed with YET is the effect of background color on accuracy. Specifically, black backgrounds led to noticeable distortions in calibration data, resulting in skewed gaze point estimations. This issue likely arose because the calibration was performed with a white background, and when the background color changes, the screen brightness changed and affected the illumination of the eye, potentially the size of the pupil and the reflection of the screen diminished on screen. This discrepancy may be particularly problematic in applications where the background of the viewed image is incongruent with the calibration background.

Another challenge occurred during eye recognition. For some users, difficulties arose while trying to recognize the eye. For most participants, the eye was not recognized immediately after the head mount was put on for the first time. The ruler sometimes had to be detached and reattached in a different angle. It appeared that a 90° angle of the participant to the center of the screen was optimal most of the time. Since a pair of headphones was used to attach the ruler with the endoscope camera to it, the ruler had to be placed diagonally in front of the face for the camera to be able to detect the eye. Sometimes, only a part of the eye was recognized which led to problems during the calibration. In these cases, the program had to be closed and started again.

Another factor influencing the accuracy and data is the participant´s efficacy in adhering to instructions, keeping their head still, and sustaining attention. The duration for each trial might have been too long with 4 to 7 minutes and the overall time of testing of around one hour might have caused some to feel understimulated. Two participants reported problems staying focused and felt tired. Another participant (P16) kept on shaking her head throughout the experiment even though she was instructed and asked to keep her head still multiple times. Small movements, shifting, shaking or tilting of the head appear to influence the scattering of the gaze points. Depending on the lighting conditions, a tilt of the head might change the illumination of the eye by affecting potential shadow from light above pointing down.

It appears that participants need to be trained to some degree to use the eye tracker since they are expected to follow instructions to keep their head still during calibration and while performing the task. Therefore, the quality and standardization of the instructions are crucial for ensuring consistent data collection across participants. For this study, participants were first asked to adjust the height of the chair, then the height of the table to find a position that was most ergonomic and could be sustained throughout each trial.

The option of adjusting the height of the table allowed for a more ergonomic position for participants. It can be expected that a more ergonomic position enables participants to keep their head still more effortlessly. Without a head adjustable table, objects such as books could be used to prop up the headrest to a position in which a participant could sit more upright rather than leaning forward. Though, this might decrease the stability of the headrest and affect the distance to the screen and tilt of the head.

The paper towel roll appeared to have sufficed for this study, but a headrest with additional support for the forehead might reduce artifacts caused by head movements, such as scattering of the gaze estimates and thereby increasing accuracy and precision.

Probable use cases

In use cases where a scenario requires users to focus on or interact with objects further than ~15° from the center of the screen this would be an issue if objects have a larger radius than the estimated 3°. Visual search tasks, navigation studies or interfaces with relevant small areas of interest at the edge of a screen would not be viable use cases.

According to the benchmarks YET could be employed for use cases demanding an accuracy of above 2.5°. Thus, use cases demanding a distinction between words or small buttons are not applicable for YET. Research investigating attention across broader areas of interest of above 2.5° would be appropriate. For instance, in web usability testing focus on general areas of interest could be studied, such as sections of a webpage like navigation menus, content blocks, or advertisements. An accuracy of 2.5° is adequate to determine whether users are engaging with larger elements, with a visual angle radius below 2.5°, like banners or general content areas, without needing precise fixation on smaller buttons or links. Given the example in Figure 2, YET should be able to discern between smaller pictures and, paragraphs and even sentences if the AOI has a lower radius than 2.5°. Further, cognitive experiments with larger areas of interest would be viable as a probable use case for YET.

Consumer research often uses eye-tracking to assess how viewers interact with advertisements. Studies focus on which parts of an advertisement attract the most attention (e.g., logos, slogans, or visual elements). Research by Wedel & Pieters (2008) demonstrates that eye-tracking can be used to study attention on large visual features, such as brand logos or product images, which are significantly larger than the YET's accuracy threshold of 2.5°.

Simple gaze-based user interaction systems in assistive technologies for disabled users where the goal is to select or differentiate between broad UI elements rather than specific buttons or fine details could be a probable use case for YET.

Eye tracking studies focusing on research about clinical characteristics or diagnosis of ADHD and ASD employ e.g face recognition tasks where particularly the focus on eyes is relevant (Levantini et al., 2020). For this probable case YET could be of use as pictures of faces could be enlarged to a size that relevant AOI of the eyes could clearly be discerned from other AOI.

Further, eye-tracking is widely used in cognitive psychology to study how people process visual information across larger areas of interest. For example, in experiments investigating how attention is distributed across broader scenes or stimuli (e.g., images, faces, or larger objects), YET would be appropriate. Research by Holmqvist et al. (2012) on eye movements during scene perception often focuses on areas much larger than 2.5°, such as studying gaze distribution on different parts of an image or identifying large objects in a visual field.

Future Directions and Research

To further enhance the feasibility and accuracy of the YET system, several areas could be further investigated.

Incongruent background

Future studies should explore whether the effect that was found for the black background occurs when the background of the presented stimulus is congruent to the calibration background. In most eye tracking applications, previewing the stimulus would affect eye tracking metrics. Thus, a blurred image could be used, or the mean screen brightness of the stimulus image could be used for the background during the initial calibration and recalibration screens.

Lighting conditions

While the lighting conditions (NIR vs LED vs noLight) did not appear to differ significantly, these lighting conditions could be further observed. It can be expected that light sources coming from the front might reduce the effect of artifacts such as small movements of the head or the head mount. The lights attached to the head mount were limited to coming from one angle, multiple LED lights from various angles might reduce the effect of shadows in the eye socket or from the eyelids. Light sources attached to the head mount might be more optimal in reducing artifacts caused by small head movements but might be more susceptible to movements of the head mount, specifically the ruler. The potential instability of the head mount might have reduced positive effects of the infrared light or integrated LED light. Another potential interaction could have been caused by other light sources. If the experiment was conducted under worse lighting conditions, maybe the lights attached to the head mount would have influenced the accuracy. Additionally, the placement of particularly the NIR might have been inadequate by being too far away or not positioned correctly. Mantiuk et al. (2012) used 3 photodiodes arranged in a triangle topology around the camera lens to assure proper illumination of the eye. Another area of research could focus on whether the effect of incongruent or dark backgrounds could be resolved with multiple lights illuminating the eye.

Comparison of different headrest designs

Future studies could explore the impact various headrest designs have on accuracy and user comfort. Comparisons between a simple headrest, a professional headrest with a forehead rest, and scenarios where the participant's head rests on a chair's headrest could help identify the optimal configuration for minimizing head movement, enhancing data quality and ensuring ergonomic testing conditions. During testing, scattering was observed, often on one axis. Specifically, for P16, who kept on shaking her head, this was noticeable. This scattering, likely caused by head movements could probably be reduced with a forehead rest. A headrest of a chair might potentially be used to stabilize the head as well while ensuring a more ergonomic position. Though, this might be more susceptible than the forehead rest to participants not adhering to instructions and intuitively moving their head.

Exploration of eye shape variability

Research focusing on different eye shapes and their effects on eye-tracking performance is essential. Since eye anatomy can vary significantly across individuals, especially concerning eye curvature, eyelid positioning, and pupil size, future studies should examine how these variations affect the accuracy and reliability of gaze detection.

Enhancement of calibration and eye recognition procedures

Investigating multiple calibration tasks and repeated eye recognition procedures could enhance the robustness of YET's calibration process. Studies could explore the feasibility of conducting multiple eye recognition tasks to maintain accuracy under slight changes in the position of the head and potentially even under different background conditions.

Implementation of multiple cameras

Figure 18 indicates an effect of accuracy on laterally asymmetric performance, with the outer left row showing larger mean Euclidian distances compared to the outer right row. The use of multiple cameras for tracking could be explored to enhance the system's accuracy and reliability. Recording both eyes might offer several potential benefits for the functionality of the YET eye-tracking system. By capturing the gaze data from both eyes, the system could utilize binocular tracking, which considers the position of both eyes to calculate a more precise point of gaze. This could reduce errors that may arise from relying on a single eye, which can be prone to deviations due to individual eye characteristics or slight head movements. Holmqvist and Andersson (2017) propose that binocularity could be used to increase eye tracking metrics, such as accuracy and precision, by calculating averages between two data samples. In addition, when tracking only one eye, any partial occlusion due to eyelids, eyelashes, or shadows can impact data accuracy. With two cameras capturing both eyes, if one eye is partially occluded, the other can still provide reliable data, thereby reducing the impact of such occlusions. This might enlarge the visual field that can be tested while maintaining good accuracy.

Limitations

While YET demonstrates potential as the world´s simplest, most affordable and accessible eye-tracking system, there are several limitations to consider. The simplicity of the system, which we tested on a minimalistic do-it-yourself head mount offers the potential for vast improvements, but at the same time comes with trade-offs.

The head mount construction, though accessible and easy to assemble, was not robust enough to prevent movement during trials. Small head movements, tilting, or shaking had an effect on the stability of the camera attached to the ruler and likely affected the precision and accuracy of the gaze estimations. A stable head mount securing the ruler more robust while making use of a counterweight should result in better performance. While the paper towel roll provided some stabilization, a more secure and ergonomic headrest with additional support for the forehead could minimize such movements and improve data precision as well.

Conclusion

The present study highlights the feasibility of YET (Your Eye Tracker) as a cost-effective, do-it-yourself alternative to traditional eye-tracking systems. While YET demonstrates promising accuracy, especially in controlled conditions such as white background settings, it faces limitations with variability across participants, inconsistencies with background brightness changes, and issues related to head movement and precision. The accuracy at the center of the screen (~2.15°) is close to established eye tracking systems, but the accuracy decreases as the gaze moves away from the center.

One key finding is that lighting conditions, whether using near-infrared, LED, or no additional lighting, did not significantly affect accuracy. However, when the background turned black, the accuracy decreased. This suggests that YET’s performance is highly dependent on maintaining consistent calibration and testing conditions, especially concerning incongruent screen brightness during testing compared to the initial calibration.

Participant variability also played a significant role in the performance of YET, with some participants showing strong accuracy and others performing poorly. Factors such as head movement stability, adherence to instructions, and physiological differences likely contributed to these discrepancies. While YET can serve as a viable solution for broader applications that do not require extreme precision, like web usability testing, general cognitive research, or consumer studies, it may not be suitable for tasks requiring high accuracy at the periphery of the screen or in cases where fine detail is essential.

In this study, YET was tested with a simplistic head mount and headrest as an affordable and accessible do-it-yourself eye tracking system enabling anyone to conduct eye tracking research. Future directions for YET could include improvements to the head mount design, different frontal lighting setups, and the potential use of multiple cameras to enhance tracking accuracy. Additionally, addressing the challenges posed by background brightness and integrating adaptive, or multiple calibration processes based on screen brightness could make YET more robust and applicable across a wider range of use cases. Overall, YET presents a promising low-cost option for eye tracking, with substantial room for optimization especially in hardware to improve performance.

References

Brand, J., Diamond, S. G., Thomas, N., & Gilbert-Diamond, D. (2021). Evaluating the data quality of the Gazepoint GP3 low-cost eye tracker when used independently by study participants. *Behavior Research Methods*, *53*, 1502-1514. https://doi.org/10.3758/s13428-020-01504-2

Belkhiria, C., Boudir, A., Hurter, C., & Peysakhovich, V. (2022). Eog-based human–computer interface: 2000–2020 review. *Sensors*, *22*(13), 4914. **https://doi.org/10.3390/s22134914**

Hansen, D. W., & Ji, Q. (2009). In the eye of the beholder: A survey of models for eyes and gaze. *IEEE transactions on pattern analysis and machine intelligence*, *32*(3), 478-500. **DOI:** 10.1109/TPAMI.2009.30

Haslwanter, T., & Clarke, A. H. (2010). Eye movement measurement: electro-oculography and video-oculography. *Handbook of Clinical Neurophysiology*, *9*, 61-79. https://doi.org/10.1016/S1567-4231(10)09005-2

Holmqvist, K., & Andersson, R. (2017). Eye tracking: A comprehensive guide to methods. *Paradigms and measures*.

Holmqvist, K., Nyström, M., & Mulvey, F. (2012, March). Eye tracker data quality: What it is and how to measure it. In *Proceedings of the symposium on eye tracking research and applications* (pp. 45-52). https://doi.org/10.1145/2168556.2168563

Houben, M. M., Goumans, J., & van der Steen, J. (2006). Recording three-dimensional eye movements: scleral search coils versus video oculography. *Investigative ophthalmology & visual science*, *47*(1), 179-187. https://doi.org/10.1167/iovs.05-0234

Housholder, A., Reaban, J., Peregrino, A., Votta, G., & Mohd, T. K. (2021, December). Evaluating accuracy of the Tobii eye tracker 5. In *International Conference on Intelligent Human Computer Interaction* (pp. 379-390). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-98404-5_36

Höffner, S. (2018). Gaze tracking using common webcams. *Institute of Cognitive Science Biologically oriented Computer Vision*.

Janthanasub, V., & Meesad, P. (2015). Evaluation of a low-cost eye tracking system for computer input. *Applied Science and Engineering Progress*, *8*(3), 185-196.

Kowalik, M. (2011). Do-it-yourself eye tracker: impact of the viewing angle on the eye tracking accuracy. *Proceedings of CESCG*, 1-7.

Levantini, V., Muratori, P., Inguaggiato, E., Masi, G., Milone, A., Valente, E., ... & Billeci, L. (2020). EYES are the window to the mind: Eye-tracking technology as a novel approach to study clinical characteristics of ADHD. *Psychiatry Research*, *290*, 113135. https://doi.org/10.1016/j.psychres.2020.113135

Li, Z., Guo, P., & Song, C. (2021, March). A review of main eye movement tracking methods. In Journal of Physics: Conference Series (Vol. 1802, No. 4, p. 042066). IOP Publishing.
DOI 10.1088/1742-6596/1802/4/042066

Mantiuk, R., Kowalik, M., Nowosielski, A., & Bazyluk, B. (2012). Do-it-yourself eye tracker: Low-cost pupil-based eye tracker for computer graphics applications. In *Advances in Multimedia Modeling: 18th International Conference, MMM 2012, Klagenfurt, Austria, January 4-6, 2012. Proceedings 18* (pp. 115-125). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-27355-1_13

Mantiuk, R. (2017). Accuracy of high-end and self-build eye-tracking systems. In *Hard and Soft Computing for Artificial Intelligence, Multimedia and Security* (pp. 216-227). Springer International Publishing. https://doi.org/10.1007/978-3-319-48429-7_20

Narcizo, F. B., de Queiroz, J. E. R., & Gomes, H. M. (2013, August). Remote eye tracking systems: technologies and applications. In *2013 26th Conference on Graphics, Patterns and Images Tutorials* (pp. 15-22). IEEE. **DOI:** 10.1109/SIBGRAPI-T.2013.8

Rakhmatulin, I. (2020). A review of the low-cost eye-tracking systems for 2010-2020. arXiv preprint arXiv:2010.05480 https://doi.org/10.48550/arXiv.2010.05480

Reingold, E. M. (2014). Eye tracking research and technology: Towards objective measurement of data quality. *Visual cognition*, *22*(3-4), 635-652. https://doi.org/10.1080/13506285.2013.876481

Shelhamer, M., & Roberts, D. C. (2010). Magnetic scleral search coil. *Handbook of Clinical Neurophysiology*, *9*, 80-87. https://doi.org/10.1016/S1567-4231(10)09006-4

da Silva Lima, D., & Ventura, D. F. (2023). A review of experimental task design in psychophysical eye tracking research. Frontiers in Human Neuroscience, 17.
doi: 10.3389/fnhum.2023.1112769

Tahri Sqalli, M., Aslonov, B., Gafurov, M., Mukhammadiev, N., & Sqalli Houssaini, Y. (2023). Eye tracking technology in medical practice: a perspective on its diverse applications. *Frontiers in Medical Technology*, *5*, 1253001.

Tatler, B. W., Wade, N. J., Kwan, H., Findlay, J. M., & Velichkovsky, B. M. (2010). Yarbus, eye movements, and vision. *i-Perception*, *1*(1), 7-27. doi: 10.1068/i0382

Todorović, D. (2006). Geometrical basis of perception of gaze direction. Vision research, 46(21), 3549-3562. https://doi.org/10.1016/j.visres.2006.04.011

Walczyk, J. J., Tcholakian, T., Igou, F., & Dixon, A. P. (2014). One hundred years of reading research: Successes and missteps of Edmund Burke Huey and other pioneers. *Reading psychology*, *35*(7), 601-621. https://doi.org/10.1080/02702711.2013.790326

Wedel, M., & Pieters, R. (2017). A review of eye-tracking research in marketing. *Review of marketing research*, 123-147.

Appendix

Tables

Table 1

Accuracy and Precision per Condition per Background

Light Condition	Background Condition	Accuracy (in °)	Precision (in °)
Infrared	White	3.88	4.17
Infrared	Black	40.6	26.3
LED	White	3.71	3.98
LED	Black	37	21.3
no Light	White	3.96	4.46
no Light	black	42.5	26.5

Table 2

Accuracy per Participant per Light Condition for the white background

Participant	IR, Accuracy (in °)	LED, Accuracy (in °)	no Light (in °)
1	6.1	5.17	5.51
2	4.08	3.88	5.74
3	3.06	3.36	4.04
4	4.48	4.92	4.27
5	5.68	4.9	6.55
6	2.55	2.63	3.14
7	2.71	2.89	2.56
8	2.86	3.15	3.24
9	3.23	2.98	3.26
10	3	3.34	3.21
11	3.19	2.95	2.92
12	3.25	3.44	3.25
13	3.67	2.76	3.17
14	3.1	2.97	3.51
15	3.54	3.75	3.65
16	6.13	6.04	3.5

17	3.35	3.54	3.34
18	3.32	3.3	3.59
19	3.46	3.03	3.51
20	3.04	2.98	3.31
21	3.36	3.12	3.16
22	3.56	3.61	3.52
23	3.66	3.18	3.54
24	8.7	7	9.61

Table 3

Accuracy and Precision per Light Condition for the white background, after filtering for the first 10 observations with change compared to without filtering

Light Condition	Accuracy (in °)	Precision (in °)	Change in Accuracy (in °)	Change in Precision (in °)
Infrared	3.09	1.49	-0.79	-1.13
LED	2.9	1.36	-0.807	-1.18
no Light	3.17	1.59	-0.793	-1.08

Table 4

Accuracy and Precision per Light Condition for the white background, after filtering for the first 10 observations and excluding the 4 Stimuli on the left side and on the right side.

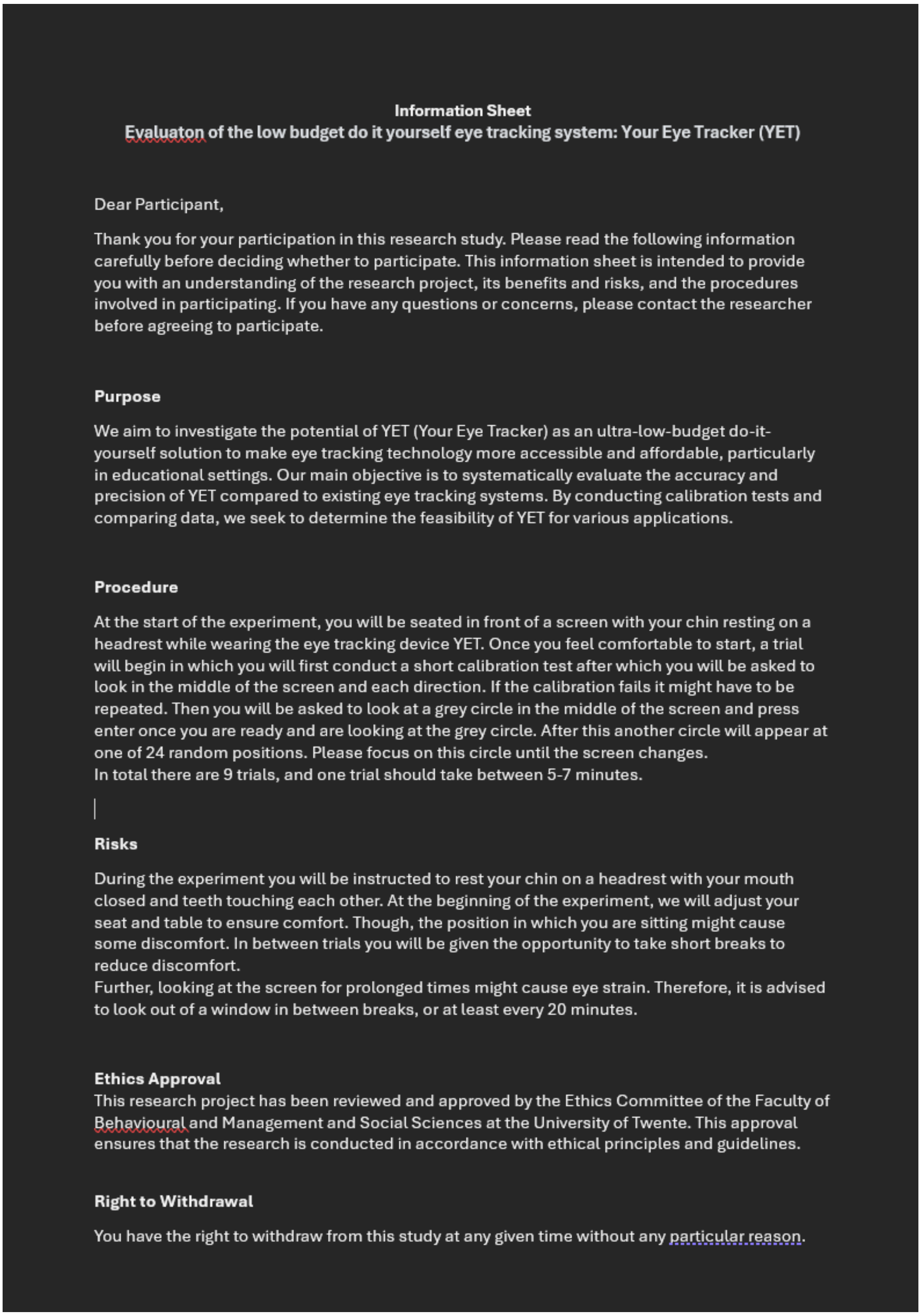
Light Condition	Accuracy excl. (in °)	Precision excl. (in °)	Accuracy orig. (in °)	Precision orig. (in °)	Change in Accuracy	Change in Precision
Infrared	2.55	3.09	3.88	4.17	-1.33	-1.08
LED	2.46	2.82	3.71	3.98	-1.25	-1.16
no Light	2.5	2.84	3.96	4.46	-1.47	-1.61

Table 5

Accuracy and Precision per Light Condition for the white background, after filtering for the first 10 observations and excluding the 4 Stimuli on the left side and on the right side.

Light Condition	Accuracy excl. (in °)	Precision excl. (in °)	Accuracy orig. (in °)	Precision orig. (in °)	Change in Accuracy	Change in Precision
Infrared	2.15	2.8	3.88	4.17	-1.73	-1.37
LED	2.14	2.84	3.71	3.98	-1.56	-1.15
no Light	2.17	2.89	3.96	4.46	-1.79	-1.56

Appendix A: Information sheet and informed consent



Information Sheet

Evaluaton of the low budget do it yourself eye tracking system: Your Eye Tracker (YET)

Dear Participant,

Thank you for your participation in this research study. Please read the following information carefully before deciding whether to participate. This information sheet is intended to provide you with an understanding of the research project, its benefits and risks, and the procedures involved in participating. If you have any questions or concerns, please contact the researcher before agreeing to participate.

Purpose

We aim to investigate the potential of YET (Your Eye Tracker) as an ultra-low-budget do-it-yourself solution to make eye tracking technology more accessible and affordable, particularly in educational settings. Our main objective is to systematically evaluate the accuracy and precision of YET compared to existing eye tracking systems. By conducting calibration tests and comparing data, we seek to determine the feasibility of YET for various applications.

Procedure

At the start of the experiment, you will be seated in front of a screen with your chin resting on a headrest while wearing the eye tracking device YET. Once you feel comfortable to start, a trial will begin in which you will first conduct a short calibration test after which you will be asked to look in the middle of the screen and each direction. If the calibration fails it might have to be repeated. Then you will be asked to look at a grey circle in the middle of the screen and press enter once you are ready and are looking at the grey circle. After this another circle will appear at one of 24 random positions. Please focus on this circle until the screen changes.
In total there are 9 trials, and one trial should take between 5-7 minutes.

Risks

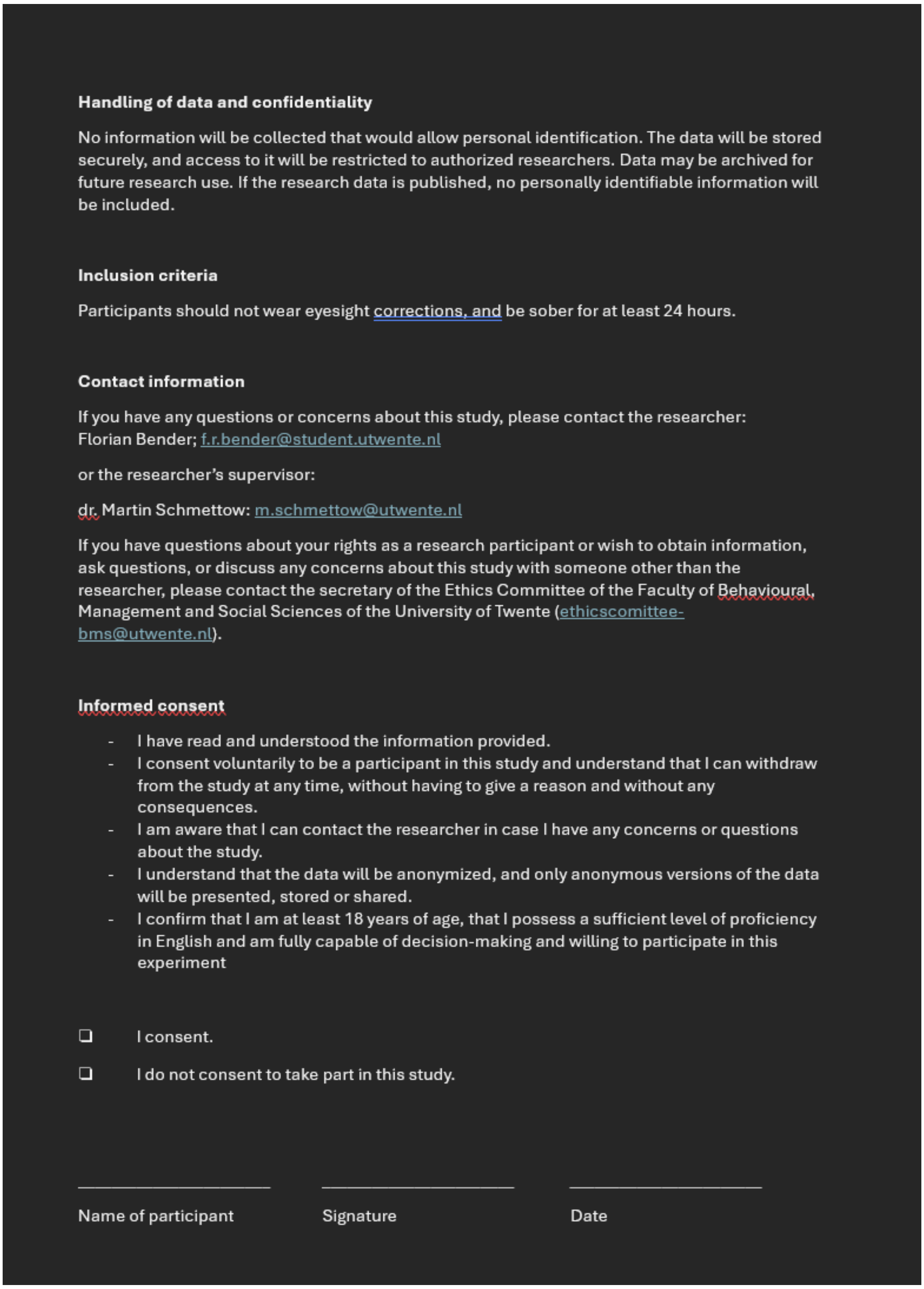
During the experiment you will be instructed to rest your chin on a headrest with your mouth closed and teeth touching each other. At the beginning of the experiment, we will adjust your seat and table to ensure comfort. Though, the position in which you are sitting might cause some discomfort. In between trials you will be given the opportunity to take short breaks to reduce discomfort.
Further, looking at the screen for prolonged times might cause eye strain. Therefore, it is advised to look out of a window in between breaks, or at least every 20 minutes.

Ethics Approval

This research project has been reviewed and approved by the Ethics Committee of the Faculty of Behavioural and Management and Social Sciences at the University of Twente. This approval ensures that the research is conducted in accordance with ethical principles and guidelines.

Right to Withdrawal

You have the right to withdraw from this study at any given time without any particular reason.



Handling of data and confidentiality

No information will be collected that would allow personal identification. The data will be stored securely, and access to it will be restricted to authorized researchers. Data may be archived for future research use. If the research data is published, no personally identifiable information will be included.

Inclusion criteria

Participants should not wear eyesight corrections, and be sober for at least 24 hours.

Contact information

If you have any questions or concerns about this study, please contact the researcher:
Florian Bender; f.r.bender@student.utwente.nl

or the researcher's supervisor:

dr. Martin Schmettow: m.schmettow@utwente.nl

If you have questions about your rights as a research participant or wish to obtain information, ask questions, or discuss any concerns about this study with someone other than the researcher, please contact the secretary of the Ethics Committee of the Faculty of Behavioural, Management and Social Sciences of the University of Twente (ethicscomittee-bms@utwente.nl).

Informed consent

- I have read and understood the information provided.
- I consent voluntarily to be a participant in this study and understand that I can withdraw from the study at any time, without having to give a reason and without any consequences.
- I am aware that I can contact the researcher in case I have any concerns or questions about the study.
- I understand that the data will be anonymized, and only anonymous versions of the data will be presented, stored or shared.
- I confirm that I am at least 18 years of age, that I possess a sufficient level of proficiency in English and am fully capable of decision-making and willing to participate in this experiment

☐ I consent.

☐ I do not consent to take part in this study.

______________________	______________________	______________________
Name of participant	Signature	Date