



# Effects of tracker location on the accuracy and precision of the Gazepoint GP3 HD for spectacle wearers

Timothy I. Murphy<sup>1</sup> · Larry A. Abel<sup>1</sup> · James A. Armitage<sup>1</sup> · Amanda G. Douglass<sup>1</sup>

Accepted: 8 November 2022 / Published online: 22 November 2022  
© The Psychonomic Society, Inc. 2022

## Abstract

Remote view eye-tracking systems are prone to errors when used on spectacle wearers due to reflections from the lenses and frame that result in inaccurate tracking. Traditionally, these trackers are situated below a computer monitor and the viewer's eye moments are recorded while they view the screen. Reflections may be influenced by the pantoscopic tilt of the spectacles, whereby the tilt angle causes incident light to be reflected to the camera. To overcome this problem, we propose mounting the tracker above the monitor to avoid these reflections and test the accuracy and precision of subjects with single vision spectacles, multifocals, and no correction, using both mounting positions. Experimental results showed that this alternate position had overall worse accuracy ( $4.06^\circ \pm 0.13$ ) and precision ( $0.67^\circ \pm 0.05$ ) compared to the standard configuration ( $2.15^\circ \pm 0.06$  vs.  $0.50^\circ \pm 0.03$ ), with more invalid readings (5.91 vs. 19.19%) for single vision lens wearers. Multifocals performed better for the top-mounting position for the top portion of the monitor, suggesting higher-order aberrations from the bottom portion of the lens negatively impact data quality. Higher pantoscopic tilt angles displayed an improved accuracy for this alternate position ( $r_{(9)} = -0.69$ ,  $p = 0.02$ ), with superior accuracy for tilt angles greater than  $14^\circ$  compared to the standard configuration. This study quantifies the impact of spectacle wear on eye-tracking performance and suggests other alternate mounting positions may be viable in certain situations.

**Keywords** Eye tracking · Spectacles · Gazepoint GP3 HD · Data quality · Glasses

## Introduction

The use of commercial eye tracking is expanding rapidly as cheaper and more user-friendly systems become available. Eye-tracking applications range from marketing insights (Pleyers & Vermeulen, 2021) to driver safety (Bitkina et al., 2021), healthcare (Murray, 2011; Murray et al., 2017; Tatham et al., 2021) and more recently in industrial and educational applications (Bitkina et al., 2021; Erkan, 2020). Two primary metrics for eye-tracking data quality are accuracy and precision. Accuracy is particularly important for studies with small areas of interest, or with very low spatial separation. Precision is important for detecting small eye movements such as fixation stability (Hunfalvay et al., 2021; Naicker et al., 2017). Spectacle wear (Carter & Luke, 2020)

and ocular misalignment or vergence posture (Poffa & Joos, 2019; Zheng et al., 2019) are reported to affect accuracy and precision in a detrimental manner. Blue eye color can also reduce precision due to the small difference in contrast between the pupil and iris as viewed by the camera (Hessels et al., 2015; Nyström et al., 2013).

Many eye-tracking systems use an integrated infrared (IR) light source and IR camera. These devices detect the IR light reflected off the cornea and at the pupil boundary to calculate eye movement using the relative positions of these images (Carter & Luke, 2020). Eye trackers can be head-mounted, with the light and camera attached to glasses or other headgear, or remote view, with the light and camera positioned distal to the participant, usually below a computer monitor and angled towards the subject's eyes. With remote view systems, spectacles can induce additional reflections from the lenses and/or frame (Carter & Luke, 2020; Gwon et al., 2014; Kübler et al., 2016; Nyström et al., 2013). These can be incorrectly interpreted as corneal reflections, resulting in erroneous gaze estimation (Huang et al., 2013). Where the IR light is reflected on to the camera

✉ Amanda G. Douglass  
amanda.douglass@deakin.edu.au

<sup>1</sup> School of Medicine (Optometry), Deakin University, 75 Pigdons Road, Waurn Ponds, Victoria 3216, Australia

and occludes the image of the pupil or cornea, gaze estimation may be negatively impacted (Gwon et al., 2014). Lens power has the potential to induce positional errors due to refraction and aberrations, with the degree of positional error proportional to lens power. (Kübler et al., 2016). Multifocal lenses also have higher-order Zernike aberrations in the lower part of the lens (Sheedy et al., 2005), which distort the image of the cornea through the spectacle lenses and initiate positional error.

Placement of a remote view tracker does not need to be limited to below the screen but should optimize the signals used by the device to detect eye positioning for the part of the screen or image desired. There are commercial systems designed to be used in different locations; for example, the Smart Eye Pro (Smart Eye AB, Gothenburg, Sweden) allows multiple cameras to be used simultaneously, with horizontal and vertical freedom of positioning. Driver monitoring systems, such as those by Seeing Machines Ltd. (Canberra, Australia), mount cameras above and/or below the driver. Given the reflection issue, systems that use the traditional placement of the camera may benefit from placement of the eye tracker above the screen by alleviating issues caused by lens reflections and higher levels of distortion viewed through the bottom of the spectacle lenses.

The Gazepoint GP3 HD is a low-cost eye tracker that has been used for a range of research projects in recent years (Karagyris et al., 2021; Sulikowski et al., 2021). It is advertised as portable, samples at 150 Hz, has a reported accuracy of 0.5–1 degree of visual angle, and is compatible with 24-inch displays or smaller – up to 25.12 degrees of visual angle from the center of the display at the ideal 65-cm working distance (Gazepoint, 2021). This has significant potential across a range of tasks, particularly in higher education. Validation data for this eye tracker exist (Brand et al., 2020; Cuve et al., 2021), however data from spectacle wearers are limited, and reflections from spectacles remain an issue (Carter & Luke, 2020). The role of pantoscopic tilt in lens reflections is well established (Fannin & Grosvenor, 1987, p. 226), and for an eye tracker placed below the screen, the angle of incidence may be close to the angle of reflection depending on this pantoscopic angle. Logically, placing the eye tracker above the monitor should reduce these reflections, and reduce distortion for multifocal wearers, however this has not been investigated. Such placement has been studied for a webcam-based eye tracker and shows similar or superior accuracy when compared to a bottom placement (Skodras et al., 2015). While reflections were not a concern for the webcam-based system as no direct light source was required, it does suggest that a top placement may be a reasonable alternative for IR systems. This paper investigates this alternate eye tracker position for the Gazepoint GP3 HD and the effects on accuracy, precision, and invalid readings for participants with and without glasses.

## Methods

This study sought to quantify the effects of alternative positioning of the Gazepoint GP3 HD above the monitor, particularly when using glasses, as this should theoretically reduce reflections. The effect on factors known to influence eye-tracking results including eye color, vergence posture, lens power, and type were also investigated in this alternative tracker position. The study was approved by the Deakin University Faculty of Health Human Ethics Advisory Group.

## Participants

Participants were recruited from the Bathurst, Geelong, and Melbourne populations in Australia through social media advertising. No specific group was targeted for this study. Those with amblyopia, strabismus, corneal pathology, or a history of ocular surgery were excluded. Vergence postures were measured using a Howell Phoria Card set (Cyclopean Design, Heathmont, Victoria, Australia) at near (33 cm) and distance (3 m), recorded as positive for exodeviation and negative for esodeviation. Eye color (blue vs. non-blue) and corrective lens strength were also recorded. Pantoscopic tilt was measured using a Shamir Pantometer (Shamir, Kibbutz Shamir, Upper Galilee, Israel).

Data from 83 tests from 41 subjects were collected. Participant attributes – eye color, lens type (no spectacles, single vision, multifocal), lens power, pantoscopic tilt, vergence postures, and history of eye conditions (astigmatism, strabismus, ptosis, anisocoria, amblyopia, nystagmus, keratoconus, dry eye) – are available in the supplementary material. One participant (single vision spectacles, blue eyes) could not complete calibration for the bottom position as the system could not detect the participant's eyes but calibrated successfully in the top position. Nine participants were unable to successfully calibrate 9/9 targets for a tracker position after four attempts, so these data were excluded from the analysis but are included in the supplementary material. From this, 72 tests from 40 subjects (33 top, 39 bottom) were used for analysis. Twenty-two of these wore spectacles (11 single vision, 11 multifocal). Pantoscopic tilt was not measured for two participants due to equipment availability.

## Experimental setup

A Philips 243B9/75 23.8-inch flat screen monitor at a native 1920 × 1080 resolution (Philips, Amsterdam, The Netherlands) was set up with two Gazepoint GP3 HD eye trackers: one mounted above the screen using an inverted Gazepoint VESA Mount (Gazepoint, Vancouver, British Columbia, Canada), and one placed below the screen



**Fig. 1** Dual Gazepoint GP3 HD setup

on a small tripod, as shown in Fig. 1. The monitor color, brightness, and contrast were calibrated using a Datacolor SpyderX Pro (Datacolor, Lawrence, NJ, USA) with default settings. Room illumination was modulated to be  $250 \pm 50$  lux from the location of the participant's eyes, with the screen off, using a Sekonic Digitalmaster L-758DR (Sekonic, North White Plains, NY, USA). Eye trackers were configured via the Gazepoint Control software v6.7.0 to run at 150 Hz and default settings. One eye tracker was used at a time, with the other switched off and its IR light power cable unplugged. Participants were situated 65 cm from the monitor with eyes aligned with the top third of the screen, per manufacturer recommendations. This created a viewable area of approximately  $45 \times 25$  degrees. No head restraints were used, and participants were requested to move their head and eyes naturally throughout the study. Eye position relative to the headbox was monitored using the Gazepoint Control software and, in the event of the participant moving outside of the bounds of the eye tracker, the test was paused and the participant moved back into position within the headbox. This occurred only once during testing, and no targets were clicked during this time.

Custom software was created to measure eye-tracker accuracy and precision, available at <https://github.com/tim-murphy/eye-tracker-validation> version 1.0.1. This software was configured to display a small target in each of the 36 locations shown in Fig. 2a. The target design was a crosshair and bullseye combination (Fig. 2b) with the diameter of the central circle set to 6 pixels (angular subtense approximately 8.28 arcmin). This design has been shown to maximize fixation stability (Thaler et al., 2013).

## Protocol

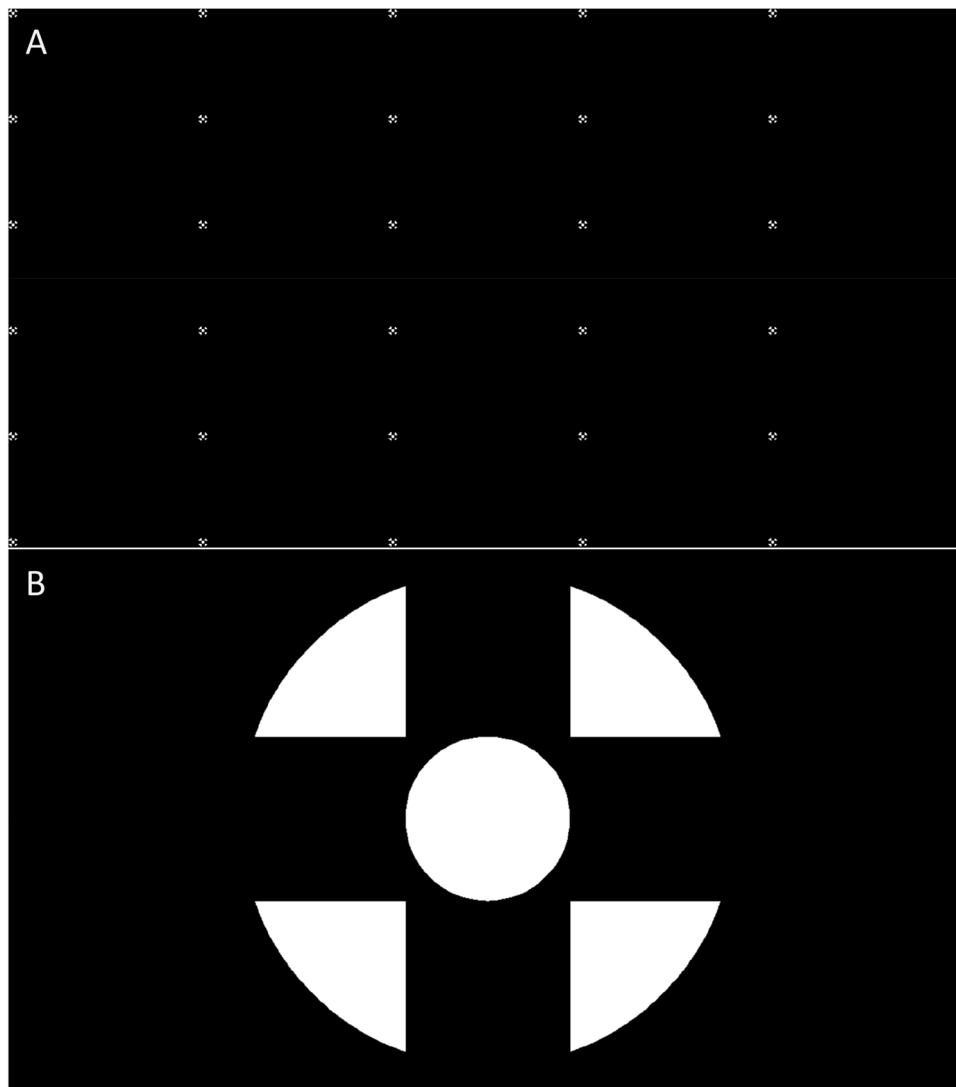
The test program was run once for each eye tracker, randomizing which tracker was used first. Before each run, the eye tracker was calibrated using a nine-point calibration, which was repeated until 9/9 targets were successfully calibrated for both eyes. Participants who failed to meet this threshold completed the task but data for this run were subsequently excluded from the results. Targets were displayed at each of the 36 locations, one at a time in random order. The user was instructed to click on the central circle on the target using a high-resolution optical mouse (Deathadder RZ01-0321, Razer Inc., Irvine, CA, USA). If the tip of the cursor was not within the central circle, this event was ignored by the software and the participant was required to continue clicking. Otherwise, the coordinates of the target, cursor, and gaze position of both eyes as measured by the eye tracker at the time of click were recorded, and the target moved to a new location.

If the gaze position was outside of the monitor limits, or the eye tracker could not calculate a gaze position at that time, the event was recorded as an invalid reading. Where a valid reading was taken for both eyes, the gaze position was calculated as the average of the two readings. This was repeated until all target locations had been presented twice (per run). As the tip of the cursor (1 pixel, approximately 1.38 arcmin) needed to align with the middle of the target (6 pixel diameter, approximately 8.28 arcmin), the resolution provided by eccentric fixation would be inadequate for this task (Kondo et al., 2008), necessitating foveation. Therefore, the gaze position at time of click was used without fixation detection or other data processing. No other eye movement metrics were calculated. This is a similar method to that used by TrackStick (Blignaut & Beelders, 2012), where a participant must click on a small target subtending  $\approx 0.1$  degrees, and the cursor position at time of click regarded as the actual gaze position. Trackstick has been validated against the Tobii T60XL (Tobii AB, Danderyd, Sweden) (Blignaut & Beelders, 2012).

At the end of the run, the participant was offered a break for as long as required. The procedure was then repeated using the other eye tracker, with the same calibration sequence taking place before testing commenced.

## Analysis

For each target, accuracy was calculated as the mean Cartesian distance from the target to the recorded gaze positions, and precision as the standard deviation of these distances (Holmqvist et al., 2011, pp. 62–64). Overall accuracy and precision were calculated in the same manner using data from all targets. A chi-squared test was used to compare the proportion of invalid readings in different tracker positions.



**Fig. 2** **a** Fixation target sizes and locations (screenshot from 23.8-inch monitor at 1920 × 1080 resolution). **b** Detail of the fixation target design

Statistical analysis was performed using the SciPy module with Python 3 (Virtanen et al., 2020). Spearman’s rank correlation was computed for each tracker position to assess the relationship between spectacle power and vergence position on accuracy and precision. A Mann–Whitney  $U$  test was used to examine the differences in accuracy and precision between subjects with and without blue eyes, and Kruskal–Wallis testing was used to compare vision correction modalities.

To determine the ideal region to use for each tracker position, K-Means clustering (Lloyd, 1982) was performed on the three-dimensional dataset of accuracy, precision, and number of readings where both eyes were invalid. The elbow method (Syakur et al., 2018) was used to determine the optimal number of clusters, with the elbow position

determined via the Kneedle algorithm (Satopaa et al., 2011). The cluster containing the point with the smallest Cartesian distance to the origin – having the highest accuracy, precision, and lowest invalid reading positions – was used to create an optimum area for target presentation.

## Results

Overall accuracy and precision were better in the bottom position, as shown in Table 1. Notably, the majority of records where the readings were invalid for both eyes were at the screen edges. The number of invalid readings was higher for the top position ( $\chi^2(2, N = 2808, M = 2376) = 681.05, p < 0.01$ ).

**Table 1** Overall accuracy and precision

Position	Accuracy	Precision	Total Readings	Invalid Both Eyes	Invalid At Edge
Top	4.06° ± 0.13	0.67° ± 0.05	2376	456 (19.19%)	341 (74.78%)
Bottom	2.15° ± 0.06	0.50° ± 0.03	2808	166 (5.91%)	144 (86.75%)

### Tracker position and factors affecting eye tracking

#### Correction type

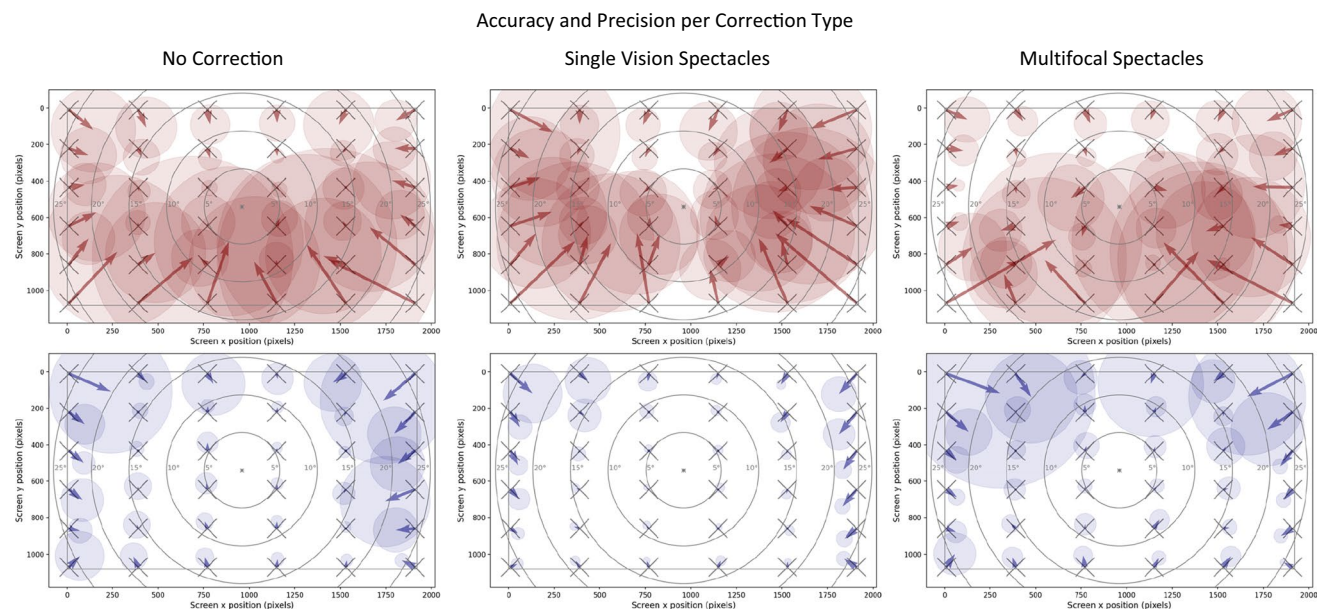
The accuracy and precision for each correction type in both top and bottom position is shown in Fig. 3, with arrows showing accuracy, and circles at the end of the arrows showing precision.

Accuracy and precision for the different correction types are shown in Fig. 4. For the bottom position, subjects without vision correction had an overall accuracy of 2.23° ± 0.09 and precision of 0.54° ± 0.06 with 56/1224 (4.58%) of readings invalid for both eyes. For single vision correction, an accuracy of 1.72° ± 0.06 and precision of 0.40° ± 0.03 was found, with 46/792 (5.81%) invalid for both eyes. The multifocal group showed an accuracy of 2.44° ± 0.14 and precision of 0.53° ± 0.06, with 64/792 (8.08%) invalid for both eyes. The effect of correction type had a significant impact on accuracy ( $H_{(2)} = 19.52, p < 0.01$ ) with Dunn testing with Bonferroni adjustment showing single vision was better than both no correction ( $p < 0.01$ ) and multifocals ( $p < 0.01$ ). There was no difference

in accuracy between multifocals and no correction ( $p = 0.70$ ). There was no difference in precision between the groups ( $H_{(2)} = 1.36, p = 0.51$ ).

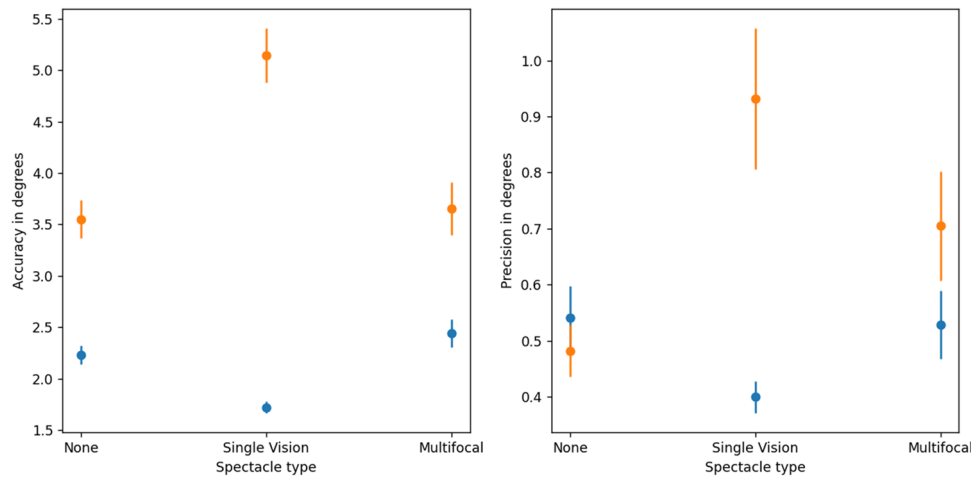
For the top position, subjects without correction had an accuracy of 3.55° ± 0.19, and precision of 0.48° ± 0.05 with 217/1080 (20.09%) invalid for both eyes. The single vision correction group gave an accuracy of 5.15° ± 0.26, precision 0.93° ± 0.13, and 132/720 (18.33%) readings invalid for both eyes. Multifocals showed an accuracy of 3.65° ± 0.25, precision 0.70° ± 0.10 and 107/576 (18.58%) readings invalid for both eyes. There was a significant difference in accuracy between groups ( $H_{(2)} = 35.13, p < 0.01$ ), with no correction better than single vision ( $p < 0.01$ ) and multifocals ( $p < 0.01$ ). There was no difference between single vision and multifocals ( $p = 1.00$ ). Precision was different between correction modalities ( $H_{(2)} = 6.00, p = 0.0498$ ), however Dunn testing showed no correction and single vision approaching significance only ( $p = 0.06$ ). No significant differences were found between the other correction types.

The number of invalid readings was worse in the top position for no correction ( $\chi^2(2, N = 1224, M = 1080) =$

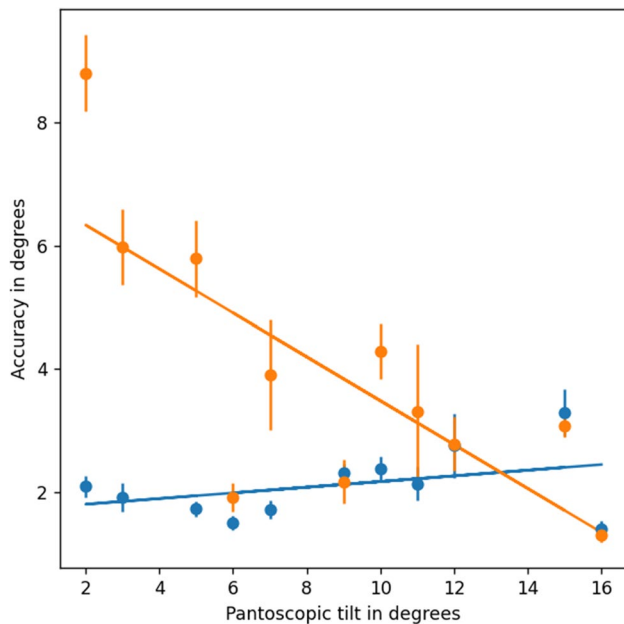


**Fig. 3** Accuracy and precision for eye-tracker positioning at the top (red) and bottom (blue) of the monitor per correction type. The black rectangle represents the edge of the monitor and grey circles indicate eccentricity isobars with the center of the screen denoted with

an asterisk. The arrows represent accuracy: the distance between the target and mean gaze position recorded by the tracker. Circles at the arrowheads represent precision: the standard deviation of the recordings



**Fig. 4** Accuracy and precision per correction type. Data for the mean accuracy and precision for the top position are shown in *orange* and bottom position in *blue*. Error bars represent SEM



**Fig. 5** Accuracy in relation to pantoscopic tilt. Data for the mean accuracy for the top position are shown in *orange* and bottom position in *blue*. The regression line is plotted as the *unbroken line* and error bars represent SEM

288.75,  $p < 0.01$ ), single vision ( $\chi^2(2, N = 792, M = 720) = 274.48, p < 0.01$ ) and multifocals ( $\chi^2(2, M = 792, N = 576) = 178.58, p < 0.01$ ).

### Pantoscopic tilt

Pantoscopic tilt had a significant effect on accuracy for the top position ( $r_{(9)} = -0.69, p = 0.02$ ) with higher angles corresponding to better accuracy as shown in Fig. 5. No

association was found for precision for the top position ( $r_{(9)} = -0.34, p = 0.31$ ), or for either statistic for the bottom position (accuracy  $r_{(9)} = 0.30, p = 0.37$ , precision  $r_{(9)} = 0.10, p = 0.77$ ).

### Spectacle power

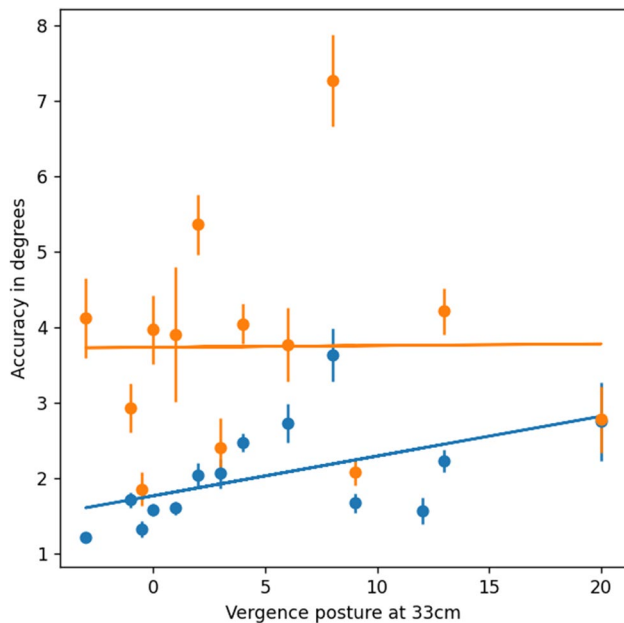
No association was found between spectacle power and accuracy for the bottom position ( $r_{(20)} = 0.27, p = 0.23$ ) or top position ( $r_{(16)} = -0.08, p = 0.76$ ), or for precision for the bottom position ( $r_{(20)} = 0.34, p = 0.12$ ) or top position ( $r_{(16)} = -0.09, p = 0.72$ ).

### Eye color

When considering eye color, subjects with blue eyes showed better accuracy (bottom:  $2.03^\circ \pm 0.09$ , top:  $3.11^\circ \pm 0.19$ ,  $U(N_{\text{bottom}} = 1218, N_{\text{top}} = 922) = 532497.5, p = 0.02$ ) and precision (bottom:  $0.46^\circ \pm 0.05$ , top:  $0.57^\circ \pm 0.07$ ,  $U(N_{\text{bottom}} = 635, N_{\text{top}} = 516) = 143802.0, p < 0.01$ ) compared to those with non-blue eyes (accuracy bottom:  $2.24^\circ \pm 0.07$  and top:  $4.94^\circ \pm 0.18$ ,  $U(N_{\text{bottom}} = 739, N_{\text{top}} = 550) = 193161.5, p = 0.06$ , precision bottom:  $0.53^\circ \pm 0.04$  and top:  $0.77^\circ \pm 0.07$ ,  $U(N_{\text{bottom}} = 739, N_{\text{top}} = 550) = 193161.5, p = 0.06$ ). There were more invalid readings for non-blue eyes compared to blue eyes ( $\chi^2(2, N = 2736, M = 2448) = 39.50, p < 0.01$ ).

### Vergence posture

Near vergence posture showed a correlation with accuracy for the bottom position ( $r_{(12)} = 0.60, p = 0.02$ ), with accuracy improving with esodeviation as shown in Fig. 6. No correlation was found for accuracy in the top position ( $r_{(12)} = 0.04, p = 0.89$ ), or for precision in either position (top:  $r_{(12)} = 0.33, p =$



**Fig. 6** Mean accuracy in relation to vergence posture for the top (orange) and bottom (blue) positions. Positive numbers indicate exodeviation and negative esodeviation. Error bars are SEM.

0.27), bottom: ( $r_{(12)} = 0.21, p = 0.47$ )). No correlations were found for distance vergence (top: accuracy ( $r_{(6)} = 0.24, p = 0.57$ ), precision ( $r_{(6)} = 0.21, p = 0.61$ ), bottom: accuracy ( $r_{(6)} = 0.12, p = 0.78$ ), precision ( $r_{(6)} = 0.40, p = 0.32$ )).

### Selecting the optimum screen area

K-Means clustering provided a subset of targets where accuracy and precision were high while invalid readings were low. The cluster containing the target with the shortest Cartesian distance to the origin was used to create an optimum area. This process was performed for all participants, shown in Fig. 7, and separately for each correction type, included in the supplementary material.

For the bottom position in this optimum area, overall accuracy was  $1.59^\circ \pm 0.03$  with a precision of  $0.37^\circ \pm 0.02$ . 51/1794 (2.84%) of readings were invalid for both eyes. For the top position, overall accuracy was  $2.63^\circ \pm 0.10$  with a precision of  $0.57^\circ \pm 0.06$  and 53/1122 (4.72%) invalid for both eyes. Accuracy was better with the bottom position  $U(N_{\text{bottom}} = 1794, N_{\text{top}} = 1122) = 766532.5, p < 0.01$ , with no difference in precision  $U(N_{\text{bottom}} = 1794, N_{\text{top}} = 1122) = 238085.5, p = 0.16$ .

### Discussion

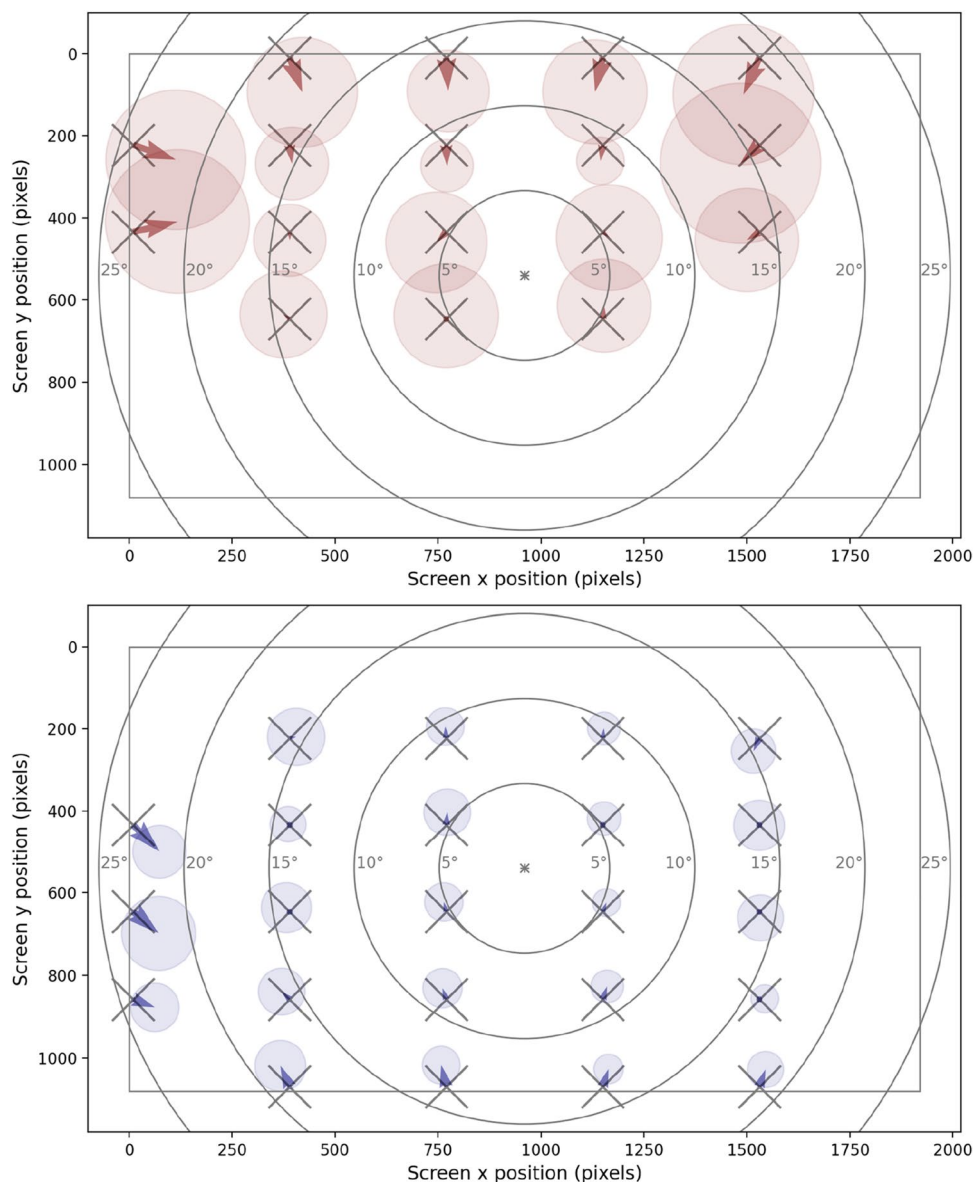
In this study, we sought to minimize the effect of reflections caused by spectacle wearers by exploring an alternative placement of the eye tracker unit, and to quantify

the effects of spectacle wear on eye-tracking performance. In theory, by moving the eye tracker above the screen (approximately  $5\text{--}10^\circ$  above the participant), lens reflections should reduce as the lenses are angled away from the camera unit. This study shows that a top-mounted unit exhibits greater validation errors and performs worse than the bottom position, with worse accuracy and precision and an increase in invalid readings. Although not formally measured, the authors anecdotally observed eyelash interference with lower targets exhibiting more invalid readings and suspect this to be a major causative factor.

This study also assessed the impact of tracker position on the accuracy, precision, and invalid reading rate with the Gazepoint GP3 HD eye tracker. The manufacturer claims a typical accuracy of  $0.5\text{--}1.0$  degrees (Gazepoint, 2021), with no published precision metrics. This has been independently validated as an accuracy of  $0.77^\circ \pm 0.70$  and precision of  $0.27^\circ \pm 0.11$  (Cuve et al., 2021), although that study did not test to the edges of the monitor and only two participants wore spectacles. With all targets considered, our data suggest significantly worse performance, with a bottom position accuracy of  $2.15^\circ \pm 0.06$  and precision of  $0.50^\circ \pm 0.03$ , and a top position accuracy of  $4.06^\circ \pm 0.13$  and precision of  $0.67^\circ \pm 0.05$ . However, the far left and right limits of tracking – greater than  $15^\circ$  from the center of the screen – exhibit a considerable degree of inaccuracy, and certainly fall outside the published values. These extremes were not included in the study by Cuve et al. (2021). Our study used the largest supported screen size and examined performance to the edges of the screen. Should a smaller screen be used, it should be placed such that it aligns with the middle region of this larger screen.

This study used an objective method for determining the optimal screen area. While previous studies have not tested to the edges of the screen area, Brand et al. (2020) found a deterioration in accuracy at the left and right side of the screen. Our results agree with these findings and suggest the best results can be obtained using the bottom mounting position, with stimuli placed in the central screen region extending to the bottom of the screen. Given the significant reduction in accuracy and precision at the left and right screen edges, these areas should be avoided unless the stimuli are sufficiently large. The optimum region calculated for the bottom position is similar to the region used by Cuve et al. (2021) to validate the same eye tracker, and the accuracy ( $1.59^\circ \pm 0.03$ ) and precision ( $0.37^\circ \pm 0.02$ ) we found agree with their findings despite the different methodologies used.

This study has explored several factors regarding spectacle performance with eye tracking. First, correction modalities exhibited different accuracy metrics. For the bottom position, single vision spectacles ( $1.72^\circ \pm 0.06$ ) performed better than no correction ( $2.15^\circ \pm 0.06$ ) or multifocals ( $2.44^\circ \pm 0.14$ ), with no difference in precision. While the cause of



**Fig. 7** Optimal screen area for the tracker placed at the top (*red*) and bottom (*blue*) of the monitor, for all participants. The *black rectangle* represents the edge of the monitor and *grey circles* indicate eccentricity isobars with the center of the screen denoted with an *asterisk*. The

*arrows* represent accuracy: the distance between the target and mean gaze position recorded by the tracker. *Circles at the arrowheads* represent precision: the standard deviation of the recordings

the superior performance for single vision spectacles is not clear, the authors suggest this could be due to the combination of lens material and antireflective coatings affecting light transmissivity, especially for shorter wavelengths (Raut et al., 2011; Shekhawat et al., 2011). The Gazepoint GP3 HD uses the PYTHON 1300 NIR semiconductor (Semiconductor Components Industries LLC, Phoenix, AZ, USA) which has a quantum efficiency curve including ultraviolet and infrared (Semiconductor Components Industries LLC, 2016, p. 8). The use of spectacles may reduce interference from

shorter wavelengths with this setup, however this warrants further investigation.

For multifocals, accuracy and precision were better with the top position for targets near the top of the screen (Fig. 3). This supports our theory of greater aberrations when the eye is viewed through the bottom portion of these lenses and suggests this alternate position could be desirable for multifocal wearers. Further, accuracy improves for the top position with pantoscopic tilt angles of 14° or greater. This suggests using an alternate position for



spectacles with a high degree of pantoscopic tilt may be advantageous.

In addition to examining tracker position, this study provides additional validation data for the Gazepoint GP3 HD. Contrary to previous studies (Hessels et al., 2015; Nyström et al., 2013), subjects with blue eyes had better accuracy and precision than those with eyes of another color, for both eye-tracker positions. There is insufficient data to determine whether spectacle correction influenced these results. Vergence posture correlated with accuracy for the bottom position only. Why no trend was seen for the top position is unclear and warrants further investigation.

Some limitations to this study need to be detailed. Positioning the participant with eyes aligned with the top one-third of the screen places the eye tracker above the eyes, potentially increasing eyelash interference. Results may vary if the top of the monitor is below the eyes. Given the high number of invalid readings for this top position, and the equal or superior performance in this top position found by Skodras et al. (2015), this warrants further investigation. Furthermore, our data do not differentiate invalid readings caused by reflections from eyelash interference or other causes. While these data are specific to the Gazepoint GP3 HD tracker, the inherent issues will be common to a range of infrared video trackers utilizing multiple camera angles.

Overall, the Gazepoint GP3 HD performs better when placed below the screen for most use cases and exhibits some variation in data quality between correction modalities. While reflections remain a problem with spectacle wear, moving the eye tracker above the monitor does not improve data quality. The exception to this may be where recording with a large pantoscopic tilt is required, or when using multifocal spectacles for targets at the top of the screen. This study tested to the edges of the largest supported screen; hence accuracy and precision metrics are not directly comparable to previous studies. Further work is required and the authors encourage validation on a wider range of eye trackers using the same software to allow a direct comparison between systems.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.3758/s13428-022-02023-y>.

## References

- Bitkina, O. V., Park, J., & Kim, H. K. (2021). The ability of eye-tracking metrics to classify and predict the perceived driving workload. *International Journal of Industrial Ergonomics*, 86, 103193.
- Blignaut, P., & Beelders, T. (2012). *TrackStick: A data quality measuring tool for Tobii eye trackers*. <https://doi.org/10.1145/2168556.2168619>
- Brand, J., Diamond, S. G., Thomas, N., & Gilbert-Diamond, D. (2020). Evaluating the data quality of the Gazepoint GP3 low-cost eye tracker when used independently by study participants. *Behavior Research Methods*. <https://doi.org/10.3758/s13428-020-01504-2>
- Carter, B. T., & Luke, S. G. (2020). Best practices in eye tracking research. *International Journal of Psychophysiology*, 155, 49–62.
- Cuve, H. C., Stojanov, J., Roberts-Gaal, X., Catmur, C., & Bird, G. (2021). Validation of Gazepoint low-cost eye-tracking and psychophysiology bundle. *Behavior Research Methods*, 1. <https://doi.org/10.3758/s13428-021-01654-x>
- Erkan, İ. (2020). Investigation of the contribution of virtual reality to architectural education. *Art, Design & Communication in Higher Education*, 19(2), 221–240.
- Fannin, T. E., & Grosvenor, T. P. (1987). *Clinical optics*. Butterworths.
- Gazepoint. (2021). *GP3 HD Eye-Tracking Device*. Retrieved January 7, 2022, from <https://www.gazept.com/product/gp3hd/>
- Gwon, S. Y., Cho, C. W., Lee, H. C., Lee, W. O., & Park, K. R. (2014). Gaze tracking system for user wearing glasses. *Sensors (Basel, Switzerland)*, 14(2), 2110–2134. <https://doi.org/10.3390/s140202110>
- Hessels, R. S., Andersson, R., Hooze, I. T. C., Nyström, M., & Kemner, C. (2015). Consequences of eye color, positioning, and head movement for eye-tracking data quality in infant research. *Infancy*, 20(6), 601–633.
- Holmqvist, K., Nyström, M., & Andersson, R. (2011). Eye Tracking. a comprehensive guide to methods and measures. In *EBL*. OUP Oxford.
- Huang, Y., Kong, W., & Li, D. (2013). Robust feature extraction for non-contact gaze tracking with eyeglasses. *Chinese Journal of Electronics*, 22, 231–236.
- Hunfalvai, M., Murray, N. P., & Carrick, F. R. (2021). Fixation stability as a biomarker for differentiating mild traumatic brain injury from age matched controls in pediatrics. *Brain Injury*, 35(2), 209–214.
- Karagyris, A., Kashyap, S., Lourentzou, I., Wu, J. T., Sharma, A., Tong, M., Abedin, S., Beymer, D., Mukherjee, V., Krupinski, E. A., & Moradi, M. (2021). Creation and validation of a chest X-ray dataset with eye-tracking and report dictation for AI development. *Scientific Data*, 8(1), 92. <https://doi.org/10.1038/s41597-021-00863-5>
- Kondo, M., Nakamizo, S., & Araragi, Y. (2008). New equally readable charts based on anisotropy of peripheral visual acuity. *Japanese Psychological Research*, 50(2), 93–99.
- Kübler, T. C., Rittig, T., Kasneci, E., Ungewiss, J., & Krauss, C. (2016). Rendering Refraction and Reflection of Eyeglasses for Synthetic Eye Tracker Images. *Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications*, 143–146. <https://doi.org/10.1145/2857491.2857494>
- Lloyd, S. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28(2), 129–137. <https://doi.org/10.1109/TIT.1982.1056489>
- Murray, I. C. (2011). *Saccadic vector optokinetic perimetry : A technique and system for automated static perimetry in children using eye tracking*. University of Edinburgh.
- Murray, I. C., Perperidis, A., Cameron, L. A., McTrusty, A. D., Brash, H. M., Tatham, A. J., Agarwal, P. K., Fleck, B. W., & Minns, R. A. (2017). Comparison of saccadic vector optokinetic perimetry and standard automated perimetry in glaucoma. Part I: Threshold values and repeatability. *Translational Vision Science & Technology*, 6(5), 3. <https://doi.org/10.1167/tvst.6.5.3>
- Naicker, P., Anoopkumar-Dukie, S., Grant, G., Modenese, L., & Kavanagh, J. (2017). Medications influencing central cholinergic pathways affect fixation stability, saccadic response time and associated eye movement dynamics during a temporally-cued visual reaction time task. *Psychopharmacology*, 234(4), 671–680.
- Nyström, M., Andersson, R., Holmqvist, K., & van de Weijer, J. (2013). The influence of calibration method and eye physiology on eyetracking data quality. *Behavior Research Methods*, 45(1), 272–288. <https://doi.org/10.3758/s13428-012-0247-4>

- Pleyers, G., & Vermeulen, N. (2021). How does interactivity of online media hamper ad effectiveness. *International Journal of Market Research*, 63(3), 335–352.
- Poffa, R., & Joos, R. (2019). The influence of vergence facility on binocular eye movements during reading. *Journal of Eye Movement Research*, 12(4). <https://doi.org/10.16910/jemr.12.4.9>
- Raut, H. K., Ganesh, V. A., Ramakrishna, S., & Nair, A. S. (2011). Anti-reflective coatings: A critical, in-depth review. *Energy and Environmental Science*, 4(10), 3779–3804. <https://doi.org/10.1039/c1ee01297e>
- Satopaa, V., Albrecht, J., Irwin, D., & Raghavan, B. (2011). Finding a “Kneedle” in a Haystack: Detecting Knee Points in System Behavior. In *2011 31st International Conference on Distributed Computing Systems Workshops, Distributed Computing Systems Workshops (ICDCSW), 2011 31st International Conference on* (pp. 166–171). <https://doi.org/10.1109/ICDCSW.2011.20>
- Semiconductor Components Industries LLC. (2016). *PYTHON 1.3/0.5/0.3 MegaPixels Global Shutter CMOS Image Sensors*. Retrieved June 15, 2022, from <https://www.onsemi.com/pdf/datasheet/noip1sn1300a-d.pdf>
- Sheedy, J. E., Campbell, C., King-Smith, E., & Hayes, J. R. (2005). Progressive Powered lenses: The Minkwitz theorem. *Optometry and Vision Science*, 82(10) Retrieved November 26, 2021, from [https://journals.lww.com/optvissci/Fulltext/2005/10000/Progressive\\_Powered\\_Lenses\\_the\\_Minkwitz\\_Theorem.14.aspx](https://journals.lww.com/optvissci/Fulltext/2005/10000/Progressive_Powered_Lenses_the_Minkwitz_Theorem.14.aspx)
- Shekhawat, N., Aggarwal, S., Sharma, A., Sharma, S., Deshpande, S., & Nair, K. (2011). Surface disordering and its correlations with properties in argon implanted CR-39 polymer. *Journal of Applied Physics*, 109, 83513. <https://doi.org/10.1063/1.3573480>
- Skodras, E., Kanas, V. G., & Fakotakis, N. (2015). On visual gaze tracking based on a single low cost camera. *Signal Processing: Image Communication*, 36, 29–42.
- Sulikowski, P., Zdziebko, T., Coussement, K., Dyczkowski, K., Kluza, K., & Sachpazidu-Wójcicka, K. (2021). Gaze and event tracking for evaluation of recommendation-driven purchase. *Sensors*, 21(4), 1381. <https://doi.org/10.3390/s21041381>
- Syakur, M. A., Khotimah, B. K., Rochman, E. M. S., & Satoto, B. D. (2018). Integration K-means clustering method and elbow method for identification of the best customer profile cluster. *IOP Conference Series: Materials Science and Engineering*, 336(1). <https://doi.org/10.1088/1757-899X/336/1/012017>
- Tatham, A. J., Murray, I. C., McTrusty, A. D., Cameron, L. A., Perperidis, A., Brash, H. M., Fleck, B. W., & Minns, R. A. (2021). A case control study examining the feasibility of using eye tracking perimetry to differentiate patients with glaucoma from healthy controls. *Scientific Reports*, 11(1), 1–10.
- Thaler, L., Schütz, A. C., Goodale, M. A., & Gegenfurtner, K. R. (2013). What is the best fixation target? The effect of target shape on stability of fixational eye movements. *Vision Research*, 76, 31–42. <https://doi.org/10.1016/j.visres.2012.10.012>
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., ... SciPy 1.0 Contributors. (2020). {SciPy} 1.0: Fundamental algorithms for scientific computing in python. *Nature Methods*, 17, 261–272. <https://doi.org/10.1038/s41592-019-0686-2>
- Zheng, X. J., Li, Z. H., Chun, X. Y., Yang, X. M., & Liu, K. (2019). A model-based method with geometric solutions for gaze correction in eye-tracking. *Mathematical Biosciences and Engineering: MBE*, 17(2), 1396–1412. <https://doi.org/10.3934/mbe.2020071>
- Open practices statement** The software used for data collection is available at <https://github.com/tim-murphy/eye-tracker-validation/tree/v1.0.1>. Participant data are provided in the supplementary material. This experiment was not pre-registered.
- Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
- Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.