Eye Tracker Data Quality Test Report

Accuracy, precision and detected gaze under optimal conditions—controlled environment

Tobii Pro Nano — Firmware: 2.33.0
Date: August 18, 2018
Metrics data analysis software version: 5.0.3

1 Introduction

This document provides an overview of the Pro Nano Eye Tracker Data Quality Tests under optimal conditions, measured in a controlled environment. The optimal conditions will serve to illustrate the best performance of the eye tracker, and how that performance varies when test conditions change during a recording.

1.1 Eye Tracker Details

Type: Screen-based/Remote eye tracker

Manufacturer: Tobii AB

Trademark: Tobii Pro

Type designation: Nano

Firmware version: 2.33.0

Serial number: TPNA1-030108014343

Sampling rate: 60 Hz

2 Method

Tests were performed in the Department of Quality Assurance test lab, at Tobii AB Headquarters, in August 2018. The lab set-up provides adequate conditions to perform data quality tests in a controlled environment. All tests were conducted by an experienced hardware technician.

2.1 Test Population

Twenty test subjects took part in the tests. The subjects were recruited from Tobii AB's local office in Stockholm, and pre-selected to exclude subjects with glasses, lenses or any type of sight correction. Table 1 displays a summary of the population characteristics.

Table 1. Describes the characteristics of the test population.

20 test participants	Composition
Age (years)	7 [21–30]; 9 [31–40]; 4 [41–50]
Phenotype	19 Caucasian; 1 Asian
Gender	4 Female; 16 Male
Eye color	6 Blue; 4 Green; 10 Brown
Eye shape	16 Round; 1 Narrow; 3 Other
Makeup	None
Sight correction	None
Eye conditions/surgery	None

2.2 Test environment and eye tracker setup

The Quality Assurance test lab is a room with no windows where lighting is provided by a LED light panel system mounted on the ceiling. The eye tracker setup is located below the panels and their orientation allows for an even distribution of light around the eye tracker and test participant. The light levels of this system can be set to different pre-defined intensities (lux units) by the test operator.

The eye tracker was placed on a plate attached to a "X-Y" rail system mounted on a table. This setup helped to change the head position of the participant relative to the eye tracker by moving the eye tracker, instead of the participant's head, to specific positions. The table is also vertically adjustable to accommodate all specified positions. A chinrest was mounted on a separate table placed directly in front of the eye tracker. Head movements were restricted to facilitate the estimation of each target's visual angle.

2.3 Task and test procedure

During each data recording session the following tests were conducted: optimal conditions—where the lighting and the position of the participants were set to obtain optimal results and matched the calibration procedure setup; different head positions in the track box—where the distance to the eye tracker, or the horizontal or vertical position of the participant relative to the eye tracker were manipulated; and different lighting environments—where the room light intensity or the stimulus background brightness were changed between tests. In total, each participant performed 10 different tests. During the data collection session, there was a short break between each test for the eye tracker operator to change the lighting or placement of the eye tracker and for the participant to rest their eyes. The order in which each test was presented, within a session, was randomized. Table 2 displays a summary of the different conditions tested.

At the start of each session, the participant was asked to sit on chair and place their chin on the chinrest. Using the "track status" information from the test software, the distance to the eye tracker was set to 65 cm. The eye tracker was calibrated once at the start of the session.

Table 2. The table describes the calibration stimulus and setup, followed by the different settings used during data collection. A single calibration was performed at the start of each session. The tests measure how robust the calibration is when faced with changes in the test environment (manipulated factor).

	Calibration	**Optimal conditions & Stimulus screen location**	**Changing lighting**	**Varying head positions**
Target number	9	12	12	12
Target screen location	Whole screen default	Whole screen	Whole screen	Whole screen
Chin rest	Yes	Yes	Yes	Yes
Lighting conditions	300 lux	300 lux	**Manipulated**	300 lux
Target/background color	White/Black	White/Black	**Manipulated**	White/Black
Placement in the track box	Center	Center	Center	**Manipulated**

The calibration procedure used was a 9-point calibration across the screen. The calibration targets were the same size as the stimulus targets, and the color was always white presented on a black background (see Figure 1). The eye tracker operator decided when to accept a calibration or perform a new one, based on the calibration feedback and own experience. In total, the operator had three attempts to improve the calibration. After which, participants were excluded from the test if no calibration data was available for one of the calibration locations.

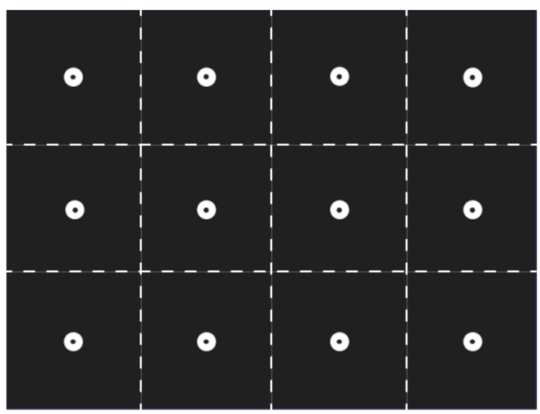


Figure 1. Illustration of the spatial location of the 12 targets during data collection. Also represented is the 3 x 4 grid used to determine the target location.

During each test, data was collected by presenting 12 white targets (with a total diameter of 36 pixels, approx. 1° visual angle) on a black background[1]. Each target has a black "bullseye" in the center. The targets were spatially arranged according to a 3 x 4 grid (individual square size of 132 x 99 mm), on a 19" LED-backlit LCD screen (900 x 1440 pixels, physical size 258 x 410 mm). Each target was in the center of each cell. The targets were presented serially (only one on the screen at a time) and in random order for 2 seconds. At the start of the session the test subjects were instructed to look at the center of the targets. Since gaze redirection takes time (saccadic latency) the first and last 500 ms of this 2-second period are discarded from the analysis. Thus, all calculations are based on a 1-second sampling window. Accuracy and precision were calculated using the center of the target as the reference.

With the exception of the values reported in section 3.3., only data from the center of the screen is used to calculate the accuracy , precision and detected gaze values for this report.

1. In the "white background" condition, the colors were reversed and targets were black with a white "bullseye" presented on a white background.

2.4 Accuracy, precision and detected gaze calculations

Accuracy was defined as the average difference between the fixation target location and the measured gaze location on the screen. Accuracy is calculated in two steps. First, by determining the mean offset between the reported gaze point and the target in millimeters, and second, by converting that value into a gaze angle based on the distance between the eye and the eye tracker (eye position). Accuracy is calculated for each eye individually and thereafter as a mean of both eyes.

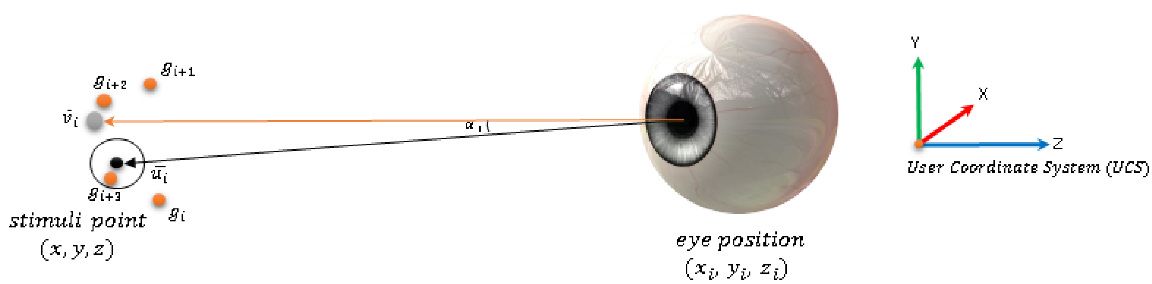
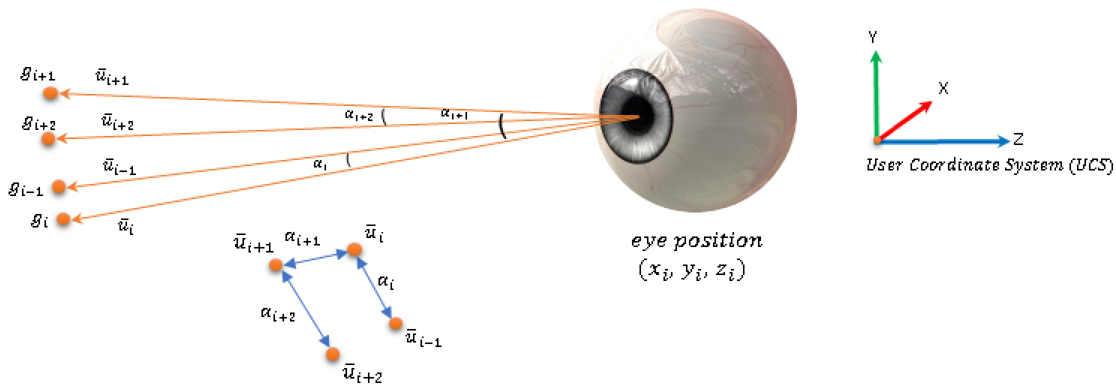


Figure 2. Illustration of the accuracy calculations. Where g_i *...* g_n *are reported gaze points,* v_i *is the average gaze point vector,* $\bar{u}_i$ *is the target point vector and* α_i *is the calculated vector angle difference between the stimuli point vector and the mean gaze point vector. Eye position and stimuli point coordinates are expressed in the User Coordinate System (UCS) provided by the Tobii Pro SDK.*

Precision is defined as the ability of the eye tracker to reliably reproduce the same gaze point measurement. Precision is calculated via the root mean square (RMS) from the successive data points (in degrees of visual angle between successive samples), both for each eye individually and as a mean from the two, i.e., it measures the variation of the recorded data via the RMS or standard variation of successive samples (SD). Precision is calculated for each eye individually and thereafter as a mean of both eyes.



$$precisionRMS_{per\ point} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \alpha_i^{\,2}}$$

Figure 3. Illustration of the precision RMS calculations. Where g_i ... g_n are successive gaze points, $\bar{u}_{i-1}$... $\bar{u}_n$ are the respective gaze vectors and α_i... α_n is the calculated vector angle difference between consecutive gaze point vectors. Eye position and stimuli point coordinates are expressed in the User Coordinate System (UCS), provided by the Tobii Pro SDK.

Detected gaze is defined as the percentage of gaze samples reported by the eye tracker during a fixation on the target. It is calculated by counting the number of invalid samples, i.e., that report that no eyes were detected, dividing that value by the total number of samples, multiplying it by 100 and finally subtracting the result from 100

$$Detected\ Gaze_{per\ point} = \left(1 - \frac{Invalid\ samples}{Total\ samples}\right) \times 100\ \%$$

All accuracy, precision and detected gaze calculations are based on raw data, collected directly from the SDK after personal calibration and subject to Two common filtering and de-noising algorithms.[2]

2. Stampe, D. M. (1993). Heuristic filtering and reliable calibration methods for video-based pupil-tracking systems. Behavior Research Methods, Instruments, & Computers, 25, 137-142.
Savitzky, A., & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. Analytical Chemistry, 36, 1627–1639.

3 Results

3.1 Summary

The average binocular accuracy, precision and detected gaze values for all tests are presented in Table 3. For head positions, the best and poorest attained value is specified for each dimension. “N” is the number of participants that met the track requirements. The track box includes offsets of 8 cm in the vertical axis, 10 cm in the horizontal axis and distances from 45 to 85 cm.

Table 3. All results. The table presents the average binocular accuracy, precision (RMS) and detected gaze results for all test conditions. A single calibration was performed at the start of each session. The tests measure how robust the calibration is when faced with changes in the test environment (e.g., varying lighting conditions). N is the number of participants who met the method requirement, i.e., used for analysis. The best and poorest value is specified for the different distances and positions measured in the head position tests. All measurements are based on raw data.

Conditions		**N**	**Accuracy (°)**	**Precision (°)**	**Detected Gaze (%)**
Optimal	Center of screen	20	0.3	0.10	99.3
	Whole screen	20	0.3	0.09	99.9
Lightning conditions (center of screen)	1 lux	20	0.8	0.09	97.8
	300 lux	20	0.3	0.10	99.3
	1000 lux	20	0.5	0.11	99.9
	White background	20	0.6	0.12	100.0
Head position (center of screen)	Z-axis	19–20	0.3–0.6	0.10–0.21	55.9–100.0
	X-axis	20	0.3–0.5	0.10–0.14	97.8–99.8
	Y-axis	20	0.3–0.8	0.10–0.11	97.0–100.0
Data processing (center of screen)	Raw	20		0.10	
	Stampe	N/A		N/A	
	Savitzky-Golay	N/A		N/A	

3.2 Accuracy, precision and detected gaze in optimal conditions

The binocular and monocular accuracy, precision and detected gaze values, under optimal conditions, are presented in Table 4. Standard deviation precision (SD Precision) is presented as a complement to the regular precision value. All participants met the calibration requirements (N=20). The average value for each metric is specified along with the standard deviation (Std).

Table 4. Accuracy, precision and percentage of detected gaze under optimal conditions. The average and monocular values are presented along with the standard deviation (Std) for each metric. Twenty *subjects were tested (N=*20*). The average ± Std SD Precision for binocular data was equal to* 0.16 ± 0.06. *Calculations are made with data from the targets located on the center of the screen.*

			Accuracy		**Precision RMS**		**Detected Gaze**	
		N	**Binocular**	**Monocular**	**Binocular**	**Monocular**	**Binocular**	**Either eye**
Optimal conditions	**Average**	20	0.3	0.3	0.10	0.14	99.3	99.5
	Std		0.2	0.1	0.05	0.04	2.6	2.0

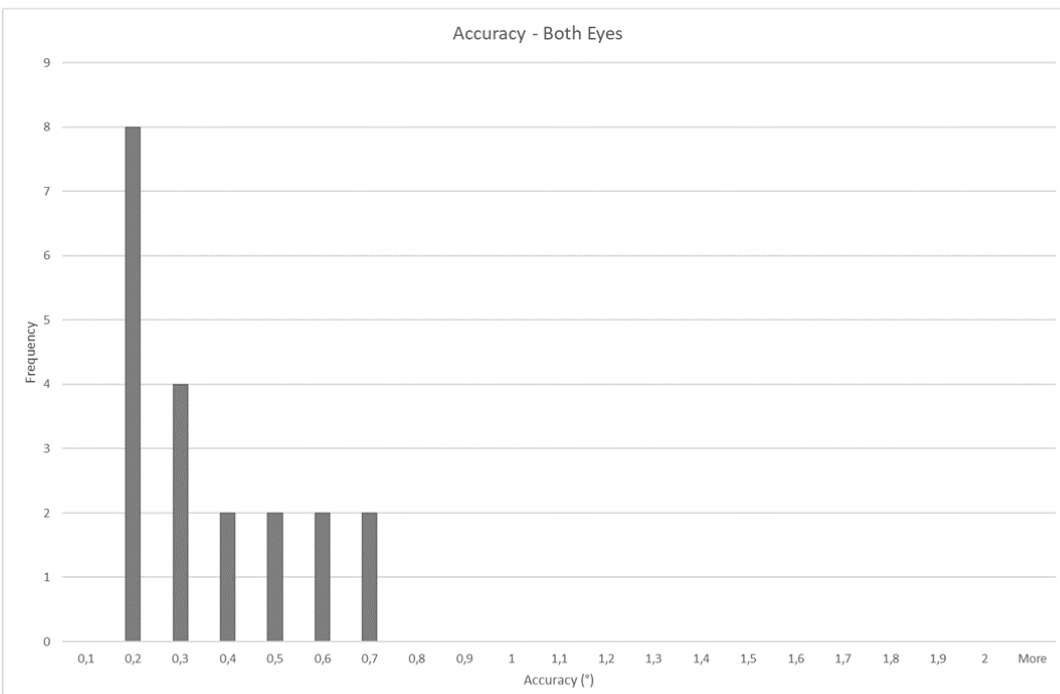


Figure 4. Average binocular accuracy distribution among the test participants for optimal conditions on the center of the screen. Results from 20 test subjects.

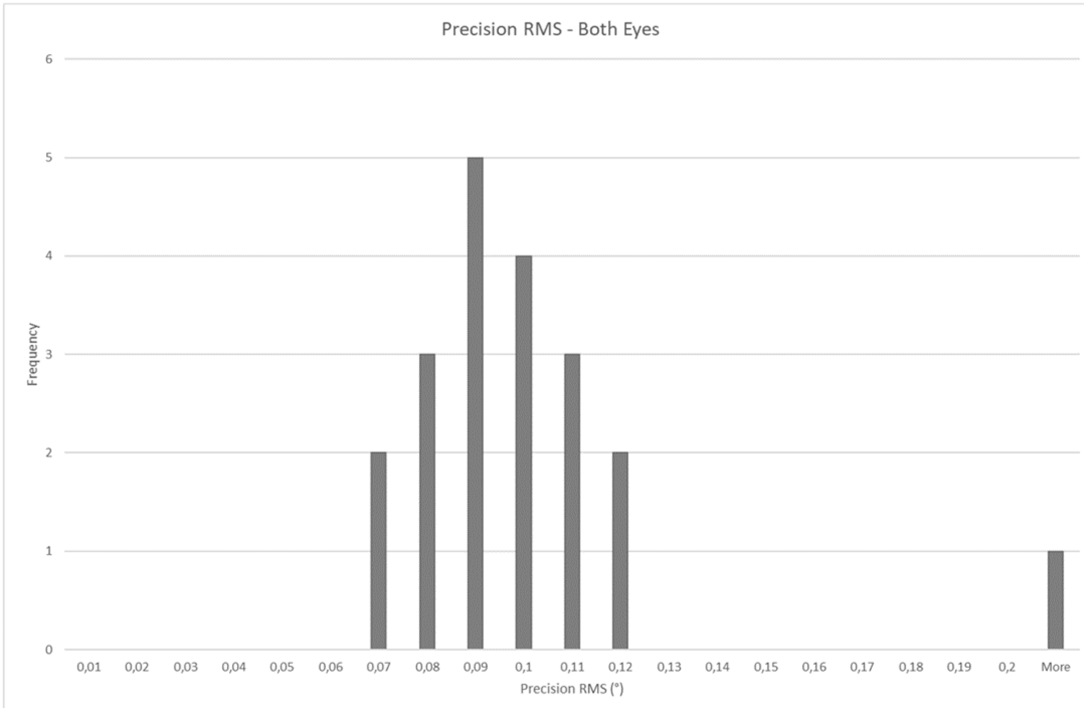


Figure 5. Average binocular precision distribution among the test participants for optimal conditions on the center of the screen. Results from 20 test subjects.

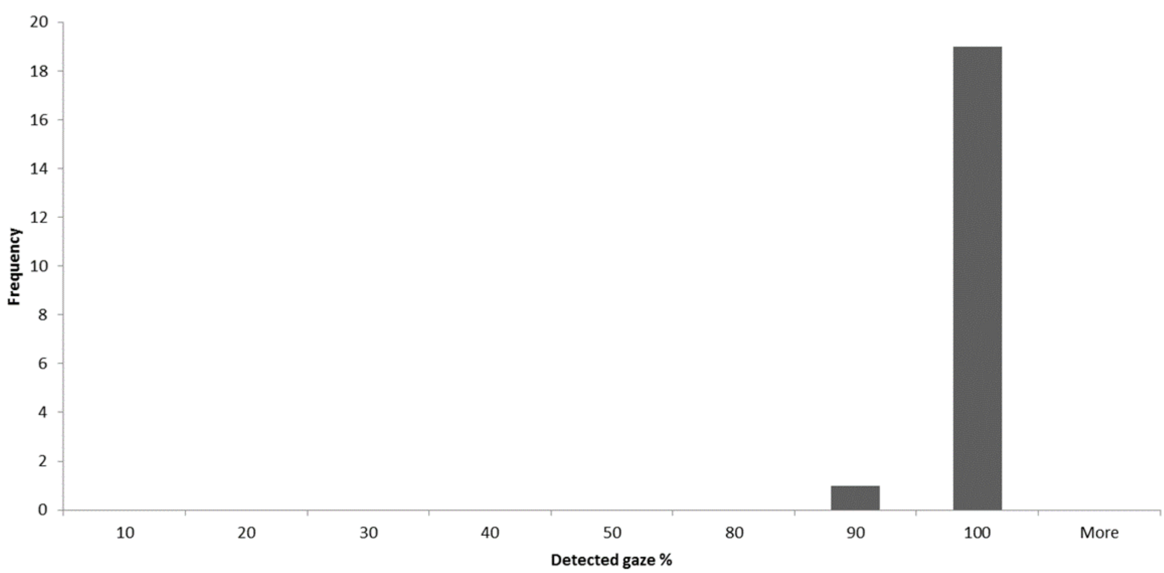


Figure 6. Average binocular detected gaze distribution among the test participants for optimal conditions at the center of the screen. Results from 20 test subjects.

3.3 Accuracy, precision and detected gaze across the whole screen

The average binocular and monocular accuracy and precision results across the whole screen are presented in Table 5. The average value for each metric is specified along with the standard deviation (Std). Figure 5 displays the accuracy, precision and detected gaze results for each target location.

Table 5. Accuracy, precision and detected gaze performance over the whole screen. Binocular and monocular values are presented for all measurements. The average values are presented along with the standard deviation (Std).

			Accuracy		Precision RMS		Detected Gaze	
		N	**Binocular**	**Monocular**	**Binocular**	**Monocular**	**Binocular**	**Either eye**
Whole Screen	**Average**	20	0.3	0.4	0.09	0.13	99.4	99.6
	Std		0.1	0.1	0.03	0.03	1.1	1.0

Accuracy: 0.51 **Precision: 0.08** **Detected gaze: 100.0** **Gaze angle: 24.9**	**Accuracy: 0.33** **Precision: 0.08** **Detected gaze: 100.0** **Gaze angle: 21.3**	**Accuracy: 0.46** **Precision: 0.10** **Detected gaze: 99.42** **Gaze angle: 21.2**	**Accuracy: 0.41** **Precision: 0.08** **Detected gaze: 99.83** **Gaze angle: 24.7**
Accuracy: 0.30 **Precision: 0.09** **Detected gaze: 98.42** **Gaze angle: 19.0**	**Accuracy: 0.34** **Precision: 0.09** **Detected gaze: 100.0** **Gaze angle: 13.6**	**Accuracy: 0.27** **Precision: 0.10** **Detected gaze: 98.58** **Gaze angle: 13.6**	**Accuracy:0.31** **Precision: 0.09** **Detected gaze: 97.67** **Gaze angle: 18.7**
Accuracy: 0.23 **Precision: 0.11** **Detected gaze: 99.50** **Gaze angle: 14.5**	**Accuracy: 0.24** **Precision: 0.10** **Detected gaze 100.0** **Gaze angle: 7.0**	**Accuracy: 0.25** **Precision: 0.09** **Detected gaze: 100.0** **Gaze angle: 7.0**	**Accuracy: 0.26** **Precision: 0.10** **Detected gaze: 99.15** **Gaze angle: 14.4**

Figure 7. Average binocular accuracy, precision and gaze detected per target spatial location on the screen. Gray cells indicate the targets used to calculate the optimal conditions values.

3.4 Accuracy, precision and detected gaze under varying lighting conditions

The binocular and monocular accuracy, precision and detected gaze results for the different lighting conditions are summarized in Table 6, as well as in Figures 8, 9 and 10. The average value for each metric is specified along with the standard deviation (Std).

Table 6. Average accuracy, precision and detected gaze under varying lighting conditions and stimuli background. The number of participants who met the tracking requirements is presented along with the binocular and monocular accuracy and precision data for each test condition.

			Accuracy		Precision RMS		Detected Gaze	
Lightning conditions		**N**	**Binocular**	**Monocular**	**Binocular**	**Monocular**	**Binocular**	**Either eye**
1 lux	**Average**	20	0.8	1.0	0.09	0.13	99.9	100.0
	Std		0.5	0.8	0.06	0.09	0.5	0.0
300 lux	**Average**	20	0.3	0.3	0.10	0.14	99.3	99.5
	Std		0.2	0.1	0.05	0.04	2.6	2.0
1000 lux	**Average**	20	0.5	0.6	0.11	0.17	100.0	100.0
	Std		0.3	0.4	0.02	0.04	0.2	0.0
White background	**Average**	20	0.6	0.7	0.12	0.18	100.0	100.0
	Std		0.4	0.5	0.03	0.4	0.0	0.0

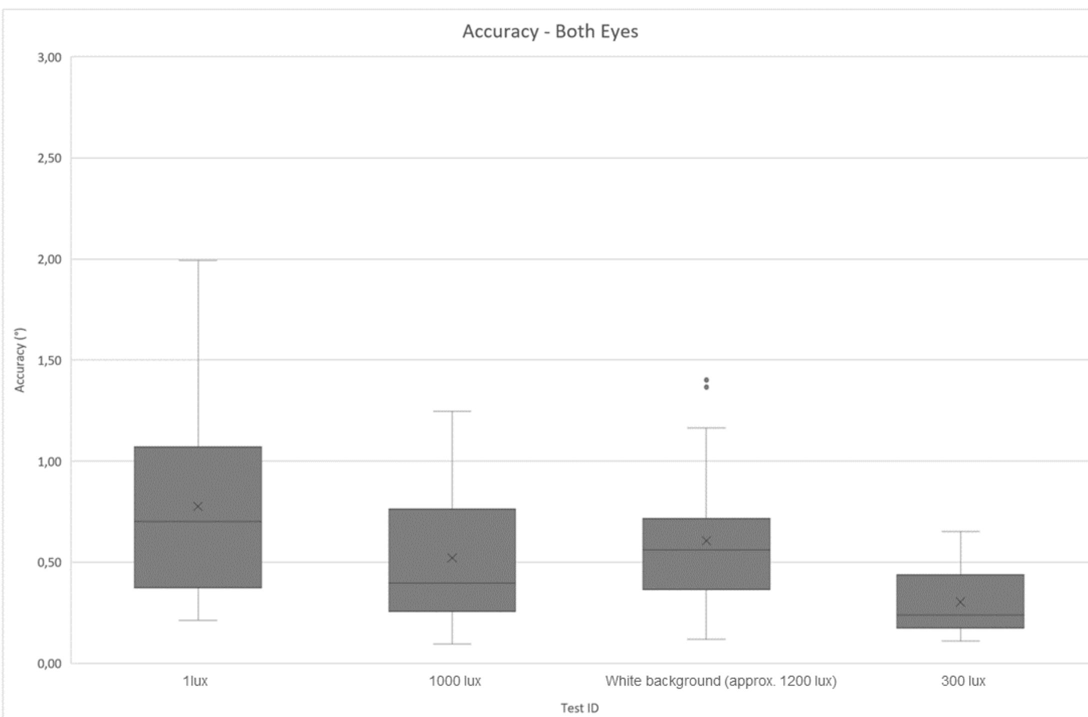


Figure 8. Binocular accuracy average values distribution for different lighting conditions. The cross in the center represents the mean; the line, the median; box edges, the quartiles; and whiskers, 3/2 times the quartile. Outliers are displayed as points. Values recorded at 300 lux represent the optimal conditions performance.

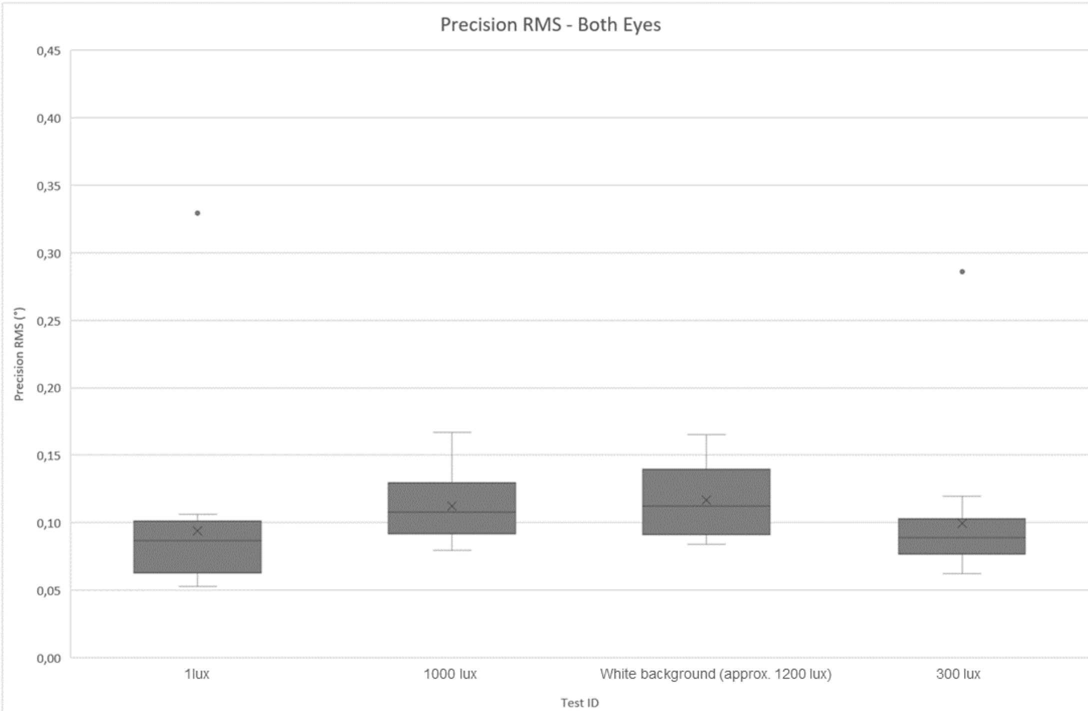


Figure 9. Binocular precision average values distribution for different lighting conditions. The cross in the center represents the mean; the line, the median; box edges, the quartiles; and whiskers, 3/2 times the quartile. Outliers are displayed as points. Values recorded at 300 lux represent the optimal conditions performance.

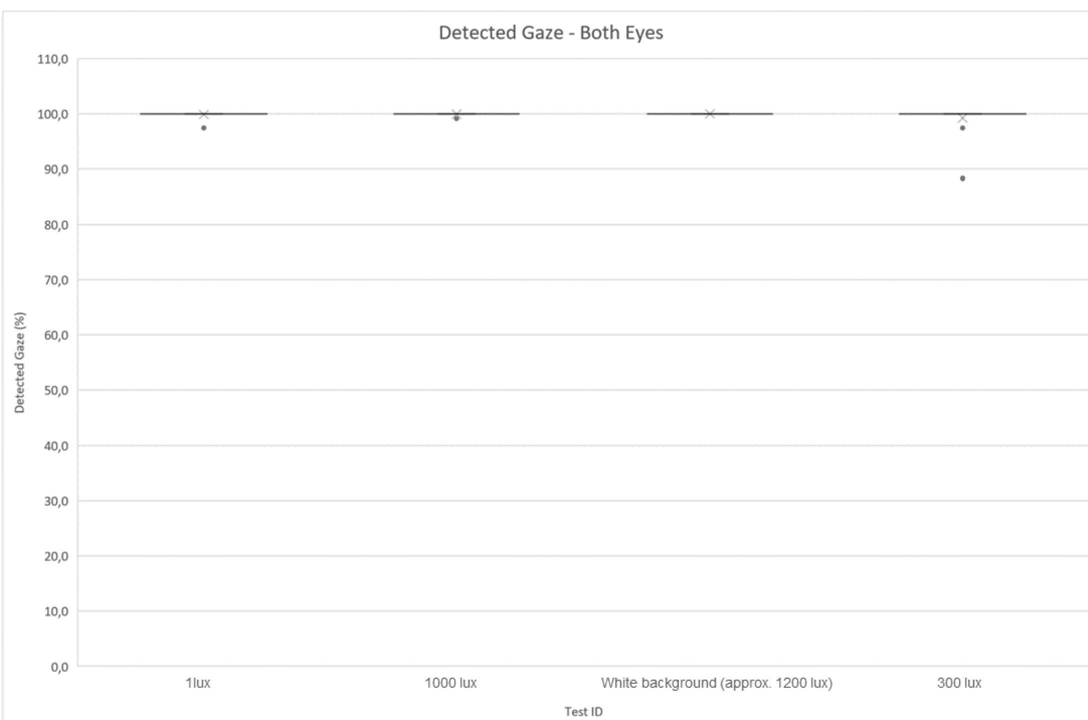


Figure 10. Binocular detected gaze values distribution for different lighting conditions. The cross in the center represents the mean; the line, the median; box edges, the quartiles; and whiskers, 3/2 times the quartile. Outliers are displayed as points. Values recorded at 300 lux represent the optimal conditions performance.

3.5 Distance from eye tracker, Z-axis

The accuracy, precision and detected gaze measured at varying distances from the eye tracker are presented in Table 7, and Figures 11, 12 and 13 respectively. The X (horizontal offset) and Y (vertical offset) positions of the test subject were kept constant, at respectively X=0.5 (track box coordinates) and Y=0.5 (track box coordinates).

Table 7, Accuracy, precision and detected gaze at varying distances from the eye tracker. The binocular and monocular accuracy and precision are presented in average values along with the standard deviation (Std) and the number of participants who met the tracking requirements (N) for each distance.

			Accuracy		**Precision RMS**		**Detected Gaze**	
Distance		**N**	**Binocular**	**Monocular**	**Binocular**	**Monocular**	**Binocular**	**Either eye**
45 cm	**Average**	19	0.6	0.8	0.19	0.29	55.9	98.5
	Std		0.2	0.3	0.07	0.11	34.5	3.1
55 cm	**Average**	20	0.4	0.4	0.10	0.14	100.0	100.0
	Std		0.2	0.2	0.02	0.03	0.18	0.0
65 cm	**Average**	20	0.3	0.3	0.10	0.14	99.3	99.5
	Std		0.1	0.1	0.05	0.04	2.6	2.0
75 cm	**Average**	20	0.5	0.5	0.11	0.17	99.8	100.0
	Std		0.3	0.3	0.02	0.04	0.6	0.0
85 cm	**Average**	20	0.5	0.6	0.21	0.34	93.7	99.8
	Std		0.2	0.3	0.14	0.31	11.1	0.9

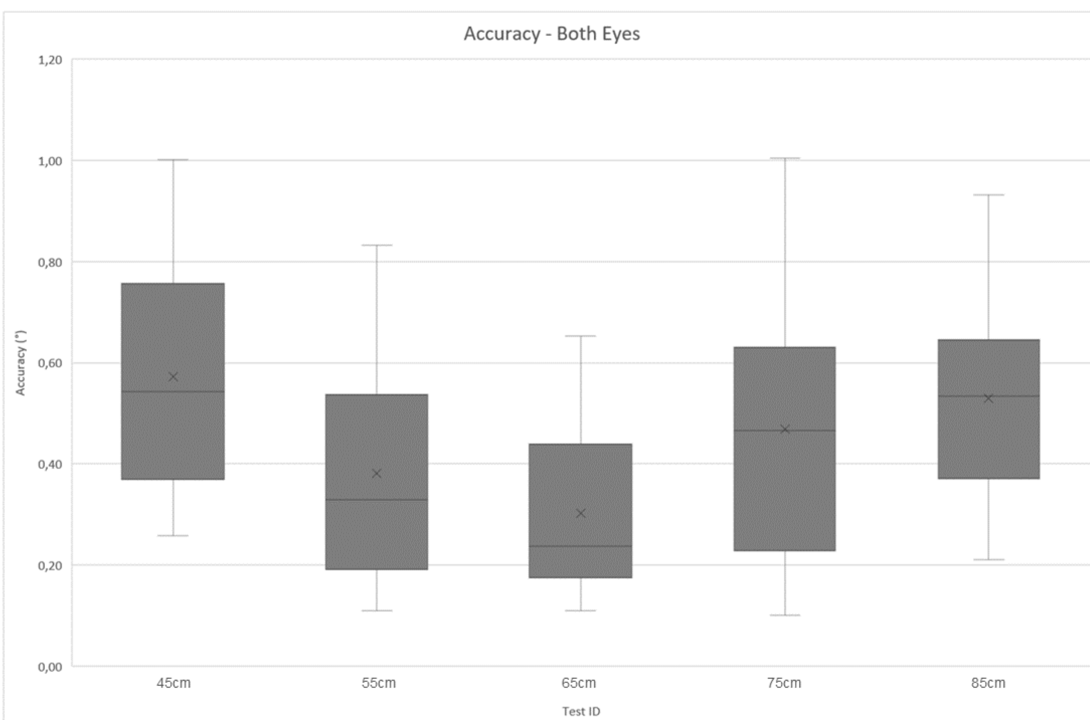


Figure 11. Binocular accuracy at varying positions in Z-axis. The cross in the center represents the mean; the line, the median; box edges, the quartiles; and whiskers, 3/2 times the quartile. Outliers are displayed as points. Values recorded at 65 cm represent the optimal conditions performance.

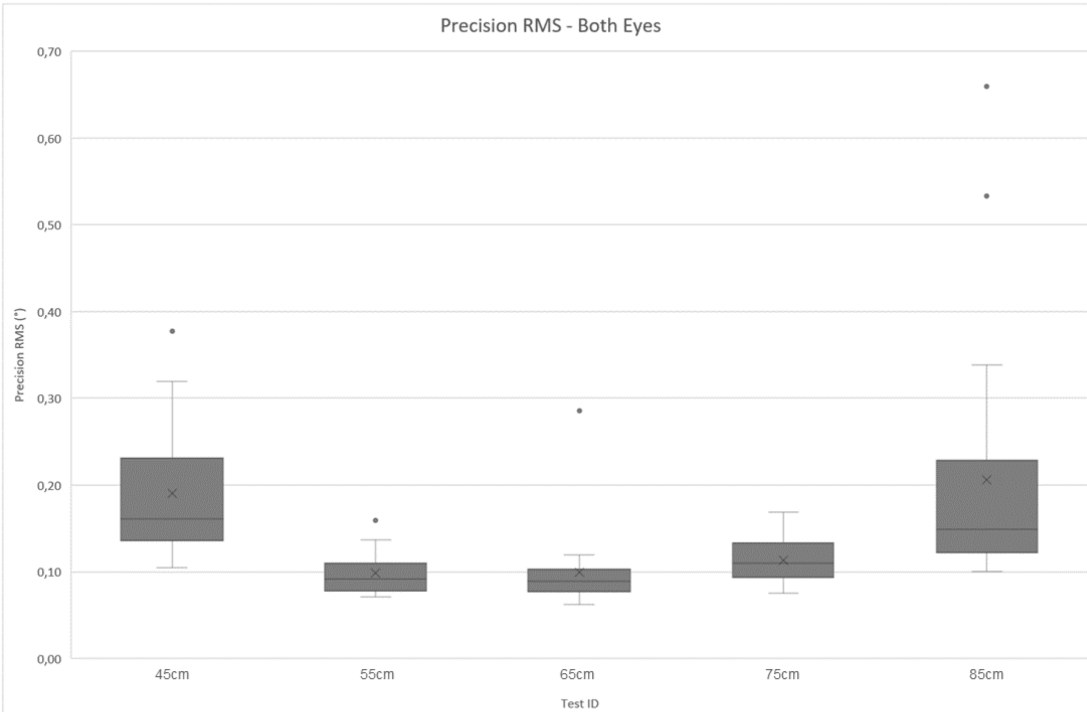


Figure 12. Binocular precision at varying positions in Z-axis. The cross in the center represents the mean; the line, the median; box edges, the quartiles; and whiskers, 3/2 times the quartile. Outliers are displayed as points. Values recorded at 65 cm represent the optimal conditions performance.

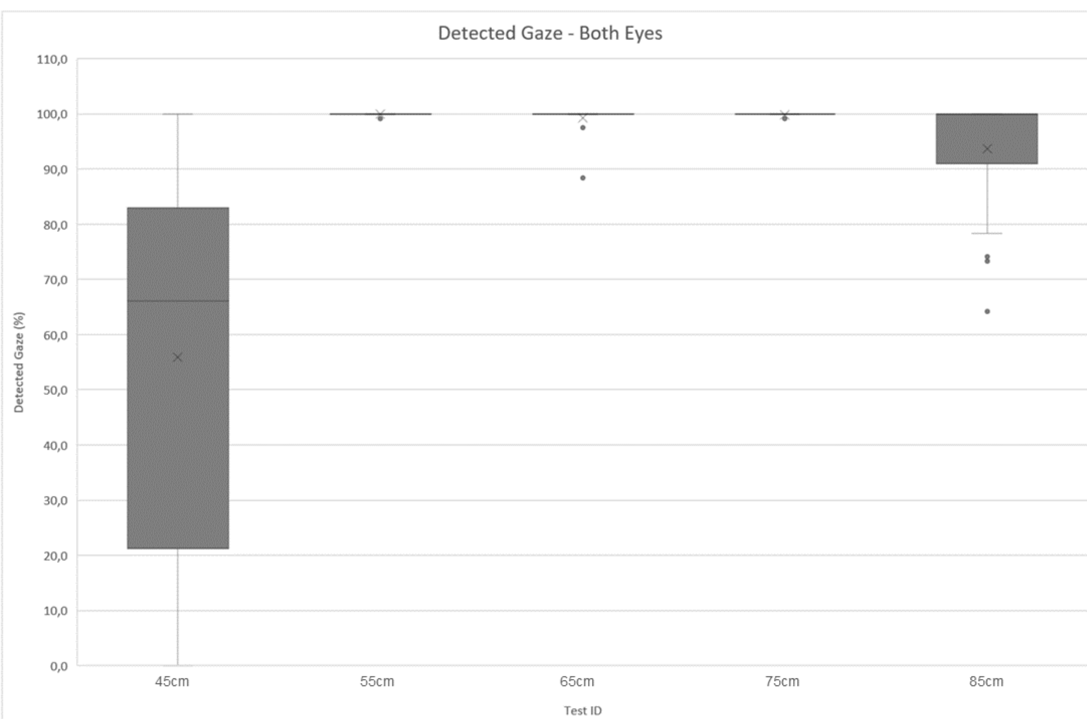


Figure 13. Binocular detected gaze at varying positions in Z-axis. The cross in the center represents the mean; the line, the median; box edges, the quartiles; and whiskers, 3/2 times the quartile. Outliers are displayed as points. Values recorded at 65 cm represent the optimal conditions performance.

3.6 Horizontal displacement, X-axis

The binocular accuracy, precision and detected gaze, measured at varying locations, horizontally from the center of the track box, are presented in Table 8 and Figures 14, 15 and 16. The Z (distance to eye tracker) and Y positions (vertical offset) of the test subject were kept constant, at respectively Z=65 cm and Y=0.5 (track box coordinates).

Table 8. Accuracy, precision and detected gaze at varying positions in X-axis. The average value for each metric is specified along with the standard deviation (Std). The number of participants who met the tracking requirements (N) is presented for each test.

			Accuracy		Precision RMS		Detected Gaze	
Distance		**N**	**Binocular**	**Monocular**	**Binocular**	**Monocular**	**Binocular**	**Either eye**
−10 cm	**Average**	20	0.3	0.4	0.10	0.13	99.8	100.0
	Std		0.1	0.2	0.02	0.03	0.9	0.2
0 cm	**Average**	20	0.3	0.3	0.10	0.14	99.3	99.5
	Std		0.2	0.1	0.05	0.04	2.6	2.0
+10 cm	**Average**	20	0.5	0.4	0.14	0.20	97.8	98.3
	Std		0.2	0.2	0.07	0.10	4.7	3.5

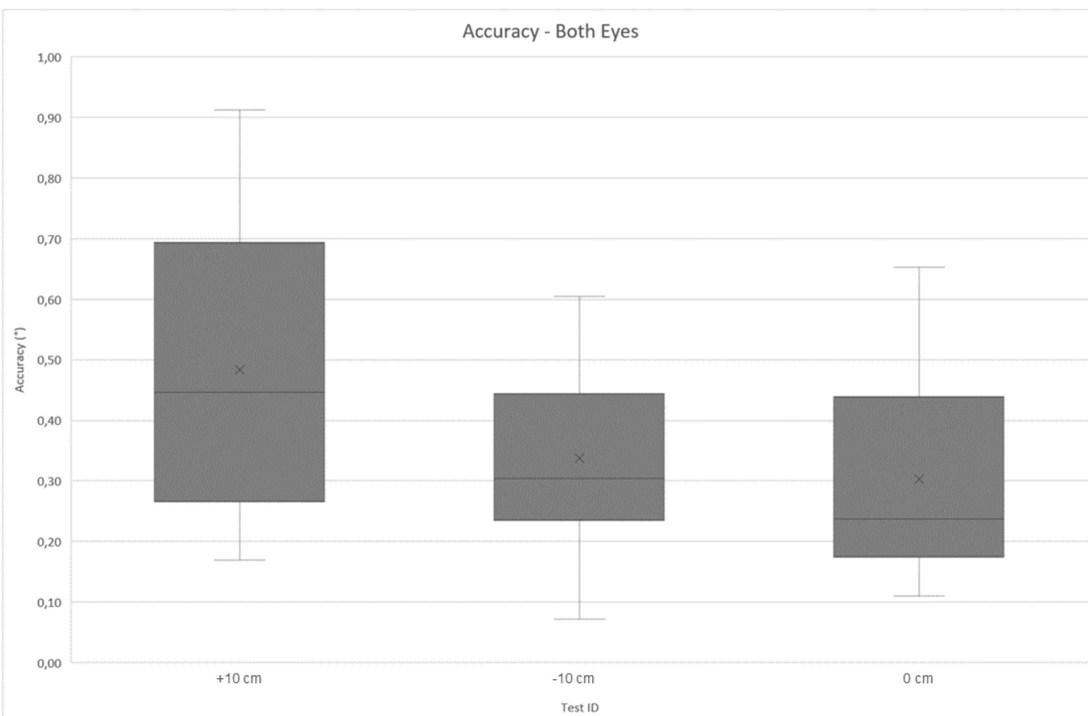


Figure 14. Binocular accuracy at varying positions in the horizontal axis (X-axis). The cross in the center represents the mean; the line, the median; box edges, the quartiles; and whiskers, 3/2 times the quartile. Outliers are displayed as points. Values recorded at 0 cm (center of the track box) represent the optimal conditions performance.

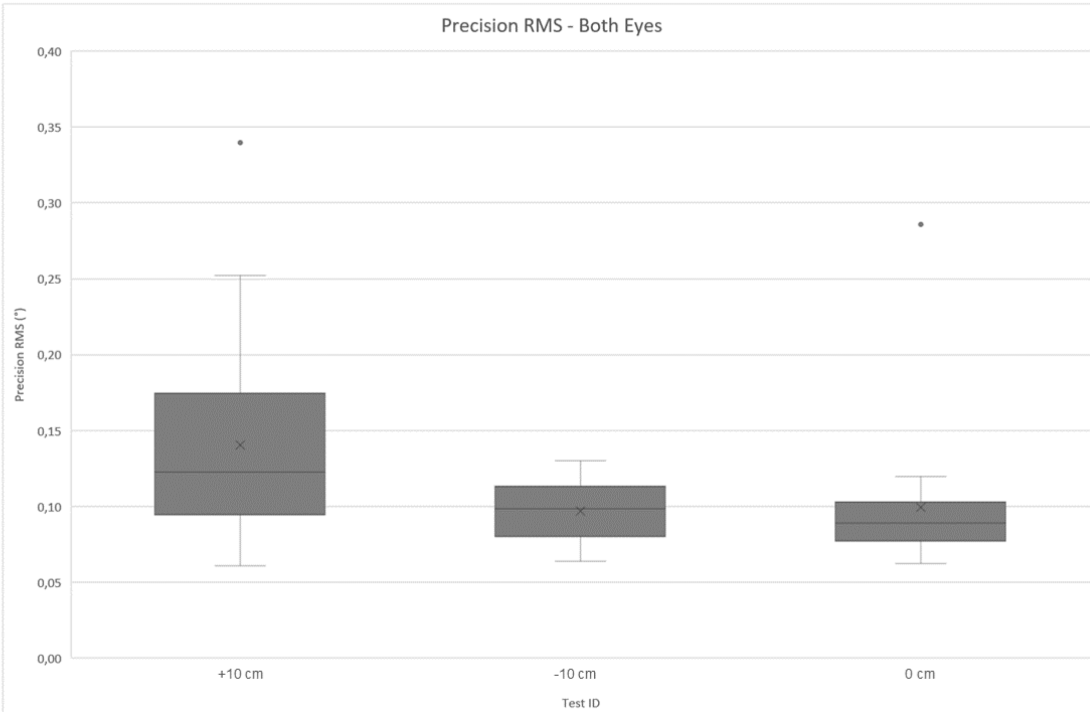


Figure 15. Binocular precision at varying positions in the horizontal axis (X-axis). The cross in the center represents the mean; the line, the median; box edges, the quartiles; and whiskers, 3/2 times the quartile. Outliers are displayed as points. Values recorded at 0 cm (center of the track box) represent the optimal conditions performance.

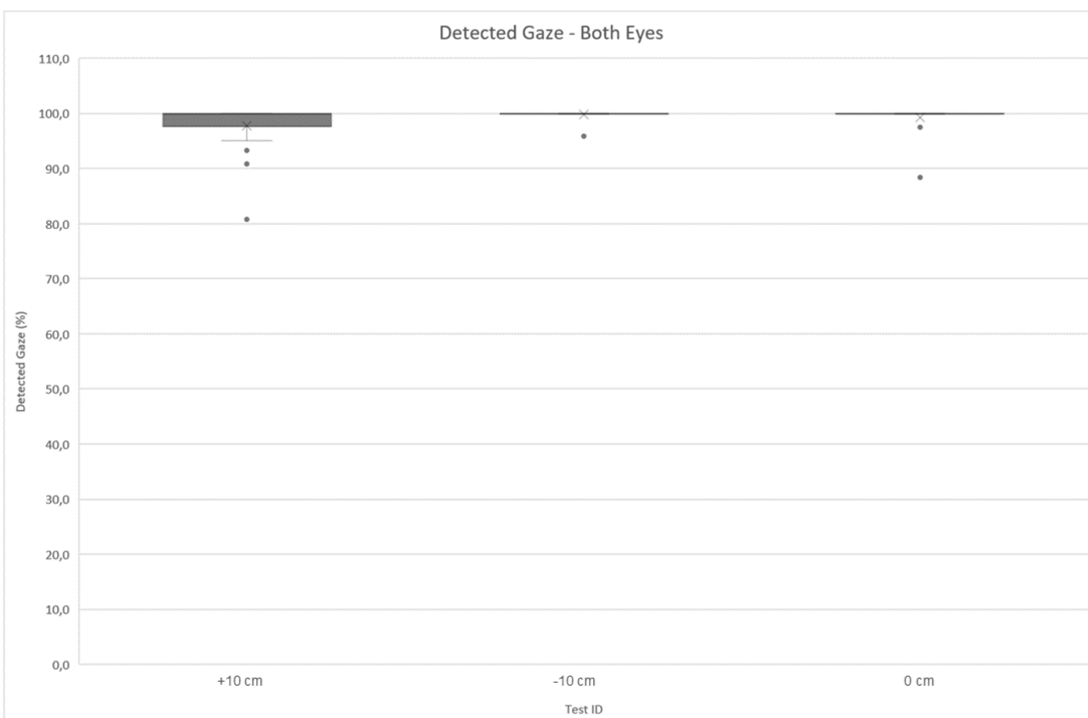


Figure 16. Binocular detected gaze at varying positions in the horizontal axis (X-axis). The cross in the center represents the mean; the line, the median; box edges, the quartiles; and whiskers, 3/2 times the quartile. Outliers are displayed as points. Values recorded at 0 cm (center of the track box) represent the optimal conditions performance.

3.7 Vertical displacement, Y-axis

The binocular and monocular accuracy, precision and detected gaze measured at varying locations, vertically from center of track box are presented in Table 9 and Figures 17, 18 and 19. The Z (distance to eye tracker) and X positions (horizontal offset) of the test subject were kept constant, at respectively Z=65 cm and X=0.5 (track box coordinates).

Table 9. Accuracy, precision and detected gaze at varying positions in Y-axis. The binocular and monocular values are presented as the average values along with the standard deviation (Std) and the number of participants who met the requirements (N) for each test trial.

			Accuracy		Precision RMS		Detected Gaze	
Distance		**N**	**Binocular**	**Monocular**	**Binocular**	**Monocular**	**Binocular**	**Either eye**
−8 cm	**Average**	20	0.4	0.5	0.10	0.15	97.0	99.6
	Std		0.2	0.3	0.04	0.06	11.0	1.8
0 cm	**Average**	20	0.3	0.3	0.10	0.14	99.3	99.5
	Std		0.2	0.1	0.05	0.04	2.6	2.0
+8 cm	**Average**	20	0.8	0.8	0.11	0.16	100.0	100.0
	Std		0.3	0.3	0.03	0.04	0.0	0.0

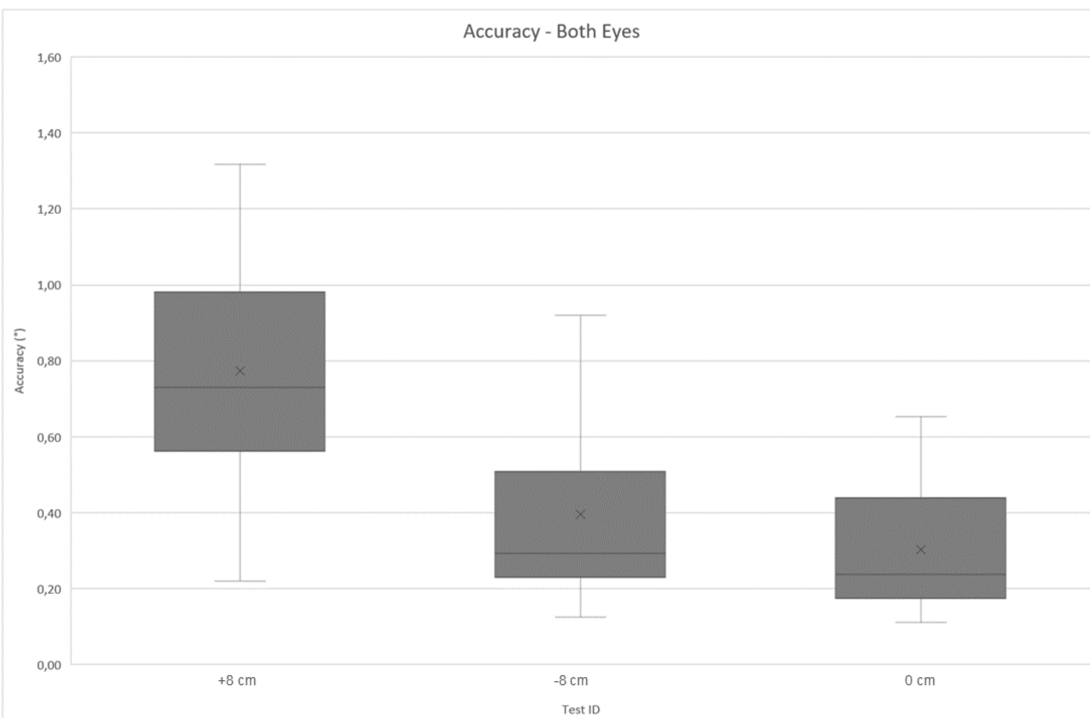


Figure 17. Binocular accuracy at varying positions in the vertical axis (Y-axis). The cross in the center represents the mean; the line, the median; box edges, the quartiles; and whiskers, 3/2 times the quartile. Outliers are displayed as points. Values recorded at 0 cm (center of the track box) represent the optimal conditions performance.

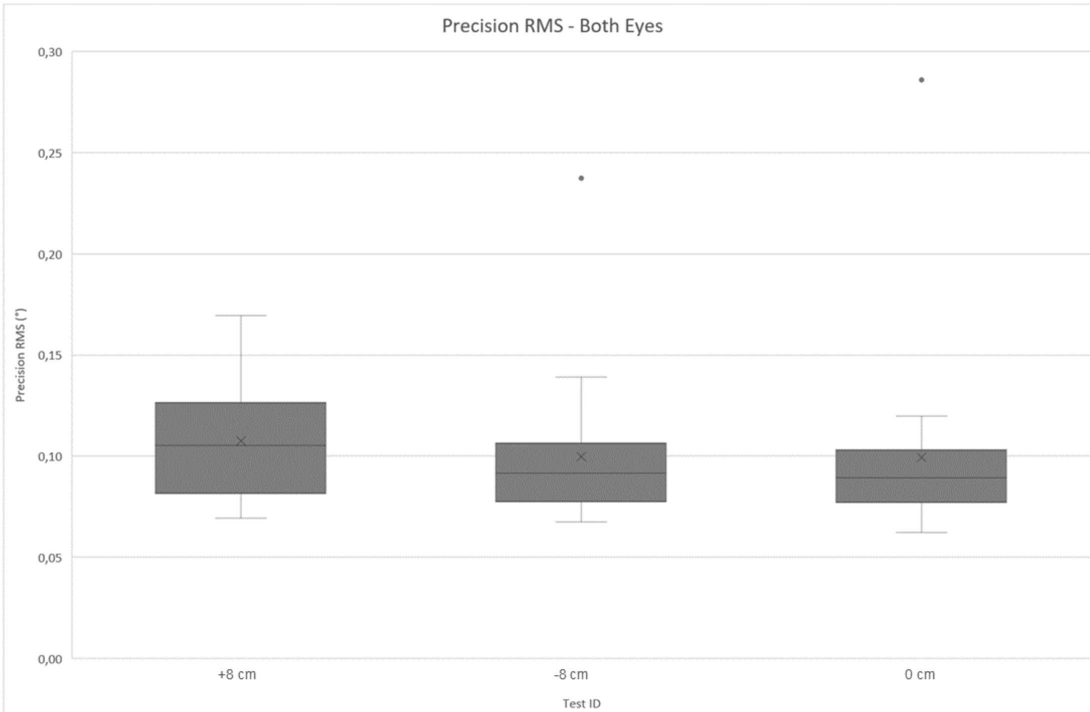


Figure 18. Binocular precision at varying positions in the vertical axis (Y-axis). The cross in the center represents the mean; the line, the median; box edges, the quartiles; and whiskers, 3/2 times the quartile. Outliers are displayed as points. Values recorded at 0 cm (center of the track box) represent the optimal conditions performance.

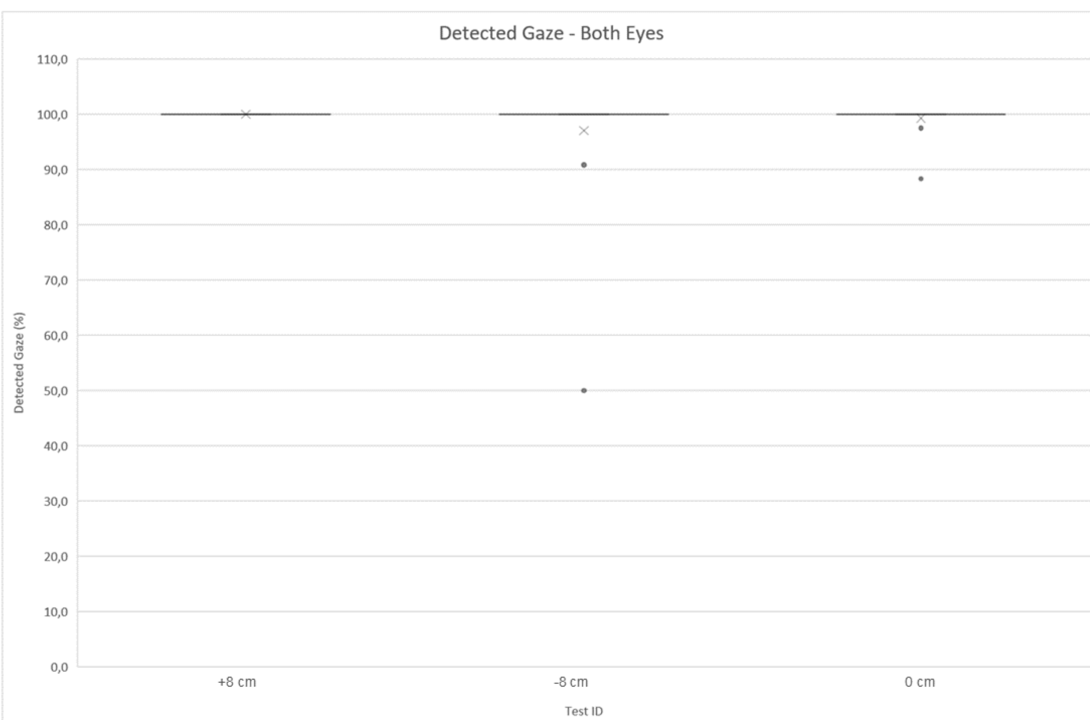


Figure 19. Binocular detected gaze at varying positions in the vertical axis (Y-axis). The cross in the center represents the mean; the line, the median; box edges, the quartiles; and whiskers, 3/2 times the quartile. Outliers are displayed as points. Values recorded at 0 cm (center of the track box) represent the optimal conditions performance.

tobii pro

Tobii Pro Support

EUROPE / GLOBAL
Phone (SWE): +46 8 522 950 10
Phone (GER): +49 69 24 75 03 4-27
support@tobii.com
Support hours: 8 am - 6 pm
Between July-August: 9 am - 5 pm
(Central European Time, GMT +1)

NORTH AMERICA
Phone: +1 703 738 1320

support.us@tobii.com
Support hours: 8 am - 8 pm
(US Eastern Standard Time, GMT -6)

JAPAN
Phone: +81 3 6420 3990

support.jp@tobii.com
Support hours: 10 am - 5.30 pm
(Japan Standard Time, GMT +9)

CHINA
Phone: +86 180 1558 5168

support.cn@tobii.com