thelightbulb

A Guide to Eye-Tracking

Table of Contents

1. Introduction	**1**
2. Background	**1**
2.1. Machine Learning	1
3. About the Algorithm	2
3.1 Processing	2
3.2 Calibration	2
3.3 Mapping	2
4. Accuracy Measurement	**3**
4.1 Defining Accuracy	3
4.2 Environment	3
4.3 User Test Flow	3
5. Mean Absolute Percentage Error (MAPE)	**4**
5.1 What is MAPE in Eye Tracking?	4
5.2 Interpreting MAPE Values	4
5.3 Importance of MAPE Eye Tracking	4
5.4 Challenges and Considerations	4
6. Accuracy Results	5
6.1 Claimed Accuracy	5
7. Evaluation Metrics for Eye Tracking Devices and Algorithms	**5**

Table of Contents (Contd.)

7.1 Angular accuracy metrics	5
7.1.1 Gaze angular accuracy	5
7.1.2 Gaze Yaw and Pitch Angular Accuracies	6
7.2 Statistical accuracy metrics	6
7.2.1 Statistical Measures of Eye Tracking Performance	6
7.2.2 Histogram-Based Metrics	6
7.3 Sensitivity metrics	6
7.3.1 Head Pose Sensitivity	6
7.3.2 Platform Orientation Sensitivity	6
7.3.3 Gaze Tracking Efficiency	7
7.3.4 Error Spatial Density	7
7.3.5 Gaze Error Analysis with Respect to Visual Eccentricity	7
7.3.6 The ROC Metric: Subjective Performance Evaluation of Eye Tracking Systems	7
8. Visualization of Eye Tracking Metrics	**7**
9. Understanding Accuracy and Errors	**9**
10. Best Practices for Improving Accuracy	**9**
10.1 Eye Tracking: Pros	10
11. Conclusion	**10**

1. Introduction

Computer vision is discipline empowering computers to glean insights from digital images and videos. It revolves around creating algorithms that emulate human perception. A pivotal application is eye tracking, a specialized field dedicated to extracting meaningful insights from visual data, particularly eye movement and behavior. This technology involves developing algorithms that decipher ocular data, mirroring human visual perception.

This technology transforms eye movements into data streams, capturing information like pupil position. By decoding eye movements, it provides valuable insights applicable across various domains. Gaze tracking, a critical application, precisely identifies and traces a person's gaze in images or videos, offering valuable data for usability studies, human-computer interaction, neuromarketing, and assistive technologies.

Our "Computer Vision" based Eye Tracking technology—simply utilize a standard webcam to precisely monitor gaze direction and duration. Enhanced by AI and Machine Learning, our system adeptly traces eye movements, specifically the iris and pupil, within the visible light spectrum, eliminating the need for specialized cameras or infrared setups.

Method Of Data Capture	Web Camera

2. Background

2.1 Machine Learning

Machine learning applied to eye tracking revolutionizes our understanding of visual attention. By training models on extensive datasets of labeled eye movement patterns, algorithms gain the ability to classify and interpret where individuals focus their attention. This technology finds applications in fields such as usability testing, market research, and human-computer interaction, enabling a deeper comprehension of user engagement and interaction nuances. Unlocking the capacity to decipher visual attention patterns allows for more informed design decisions, personalized user experiences, and heightened insights into human-machine interactions.

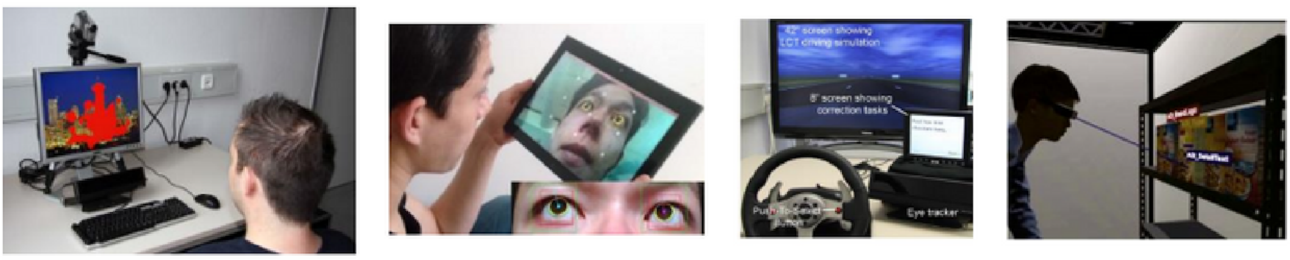


Fig. 2.1 Eye gaze applications in various consumer platforms (from left:
(a) desktop (b) tablet (c) automotive (d) head-mounted (eye tracking setups).

3. About the Algorithm

The pipeline in eye-tracking is comprised of three algorithms: **processing, calibration, and mapping,** all of which use deep learning networks to estimate the gaze from the recording. All steps take place on webcam videos captured during stimulus presentation.

3.1 Processing

Processing extracts the respondent's gaze position expressed in the camera coordinate system, relying on deep learning models to estimate the gaze position from the webcam video. The models used in this processing step are pre-trained on thousands of faces using deep learning algorithms.

3.2 Calibration

Calibration is the process of estimating the geometric properties of a participant's eyes as the basis for a fully customized and accurate gaze point calculation.

It is based on the recordings when the respondent was gazing at specific calibration markers. To calibrate the algorithm, studies are performed by positioning the respondents in front of the computer screen configured with eye tracking software while the respondent is required to keep their head position fixed, although they can move a limited amount, as long as it is within the limits of the eye tracker's range. This range is called the headbox.

During the experiments, an UI is run on the desktop or tablet screen and users are asked to follow the moving dot as it moves. This ensures that the user's fixation distance is closest to stimuli locations. The eye tracker calibration uses calibration points and the calibration stimulus comprises of dots appearing at the corners, top and bottom locations of the display.

After the calibration procedure, the calibration quality is validated by the eye tracker software, and for poor calibration, the process is repeated. This ensures that the user's fixation distance is closest to stimuli locations. The locations traced on-screen positions are recorded in pixel coordinates. The collected data comprises of the respondent's gaze coordinates on the screen and corresponding time stamps as estimated by the eye tracking software and this data is stored.

Fig 3.1 Calibration

3.3 Mapping

The model determined in the previous calibration step is then used to map the gaze position on the screen in raw (x, y) coordinates on each stimulus frame, based on the gaze data identified in the processing step. Outcome of the study is recorded as gaze dataset which is then analyzed and used for development of metrics and visualizations by the eye tracker. After completion of each study, the steps are repeated for every respondent.

**Source: sensors - Performance Evaluation Strategies for Eye Gaze Estimation 2018

The eye tracker coordinate system has its origin at the center of the frontal surface of the eye tracker which is aligned with the center of the screen. The tracker x-axis points horizontally towards the user's right, the y-axis points vertically towards the user's up and the z-axis points towards the user, perpendicular to the front surface of the eye tracker. The gaze data comprise of eye locations of a user tracked by the eye tracker and mapped into the 2D coordinates of the display screen. The gaze x, y data of user eye locations using this coordinate system has (0,0) at the display screen center and z data represents the user distance from the tracker starting from 0 at the tracker.

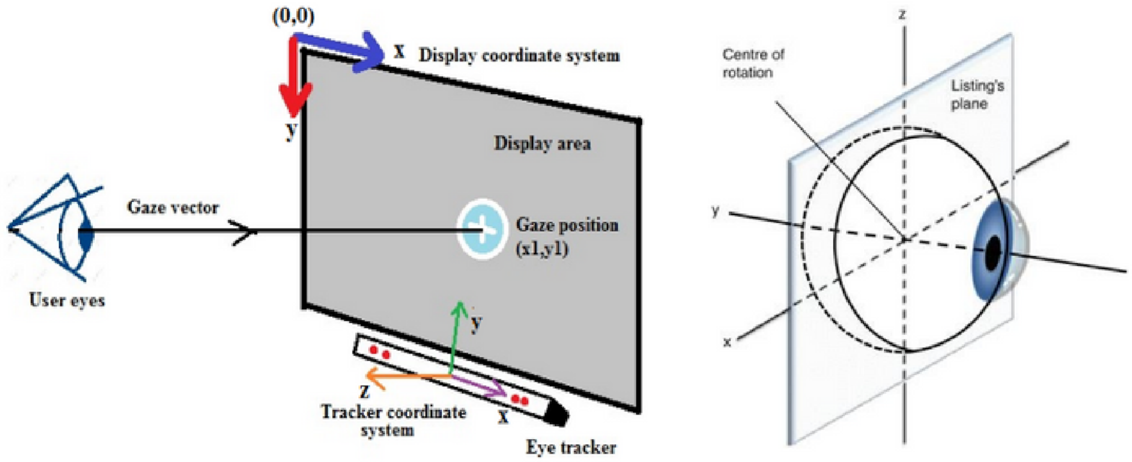


Fig 3.2 (left) Shows the display and eye tracker coordinate systems for this work and (right) shows the eye movement coordinate systems with Listing's plane and Fick's axes.

4. Accuracy Measurement

4.1 Defining Accuracy

Accuracy is used as an indicator of the eye tracker data validity. It is defined as the average difference between the real stimuli position and the measured gaze position. It is the offset between the true gaze position and the recorded position.

4.2 Environment

We have conducted the study in a controlled environment and wild environments.
The **"Controlled"** environment stimuli were displayed on a 15' laptop computer. The respondents sat 65cm-90cm away from the laptop in a well-lit room. The eye tracking data was collected using the laptop webcam with a resolution of 1080x1920 and a sampling rate of 30fps. We had certain parameters controlled, specifically the tester being seated at a desk without any glasses under proper lighting while taking the test.

Aside from the ideal conditions, additional conditions were tested: participants wearing glasses, a low webcam resolution and participants moving their heads to a certain limit. Environment may or may not vary in one or more settings mentioned above during the test.

4.3 User Test Flow

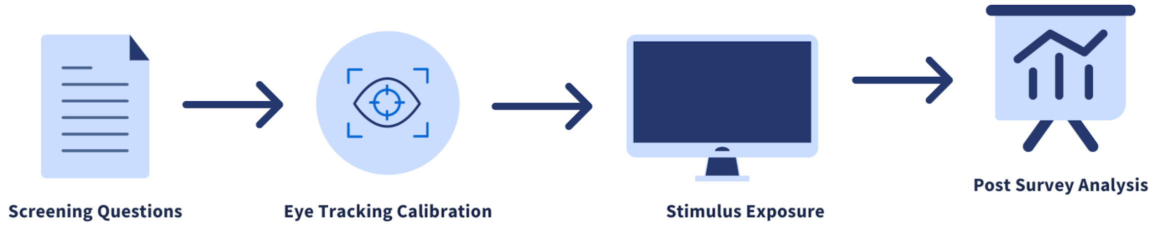


Fig 4.1 User test flow

5. Mean Absolute Percentage Error (MAPE)

Understanding and Utilizing MAPE (Mean Absolute Percentage Error) Parameters in Eye Tracking
Eye tracking technology is an invaluable tool in various fields, from market research to human-computer interaction. It allows us to comprehend and quantify visual attention, providing crucial insights into user behavior. As we delve into the metrics that gauge the accuracy of eye tracking systems, one parameter stands out — Mean Absolute Percentage Error (MAPE).

5.1 What is MAPE in Eye Tracking?

MAPE is a metric used to evaluate the accuracy of eye tracking systems by quantifying the disparity between predicted gaze points and actual gaze points. It provides a percentage-based measure of the average error in predicting where a person is looking. In essence, MAPE offers a comprehensive understanding of how closely the eye tracking system aligns with the real gaze behavior.

5.2 Interpreting MAPE Values

Low MAPE (Closer to 0%): Indicates a high level of accuracy, suggesting that the predicted gaze points closely match the actual gaze points. This is ideal for reliable eye tracking systems.

High MAPE (Closer to 100%): Indicates a larger discrepancy between predicted and actual gaze points, signifying lower accuracy. Systems with high MAPE values may need refinement for improved precision.

5.3 Importance of MAPE Eye Tracking

Accuracy Assessment: MAPE is a pivotal tool for assessing the accuracy of eye tracking systems. It offers a quantitative measure that goes beyond qualitative observations, enabling a more nuanced understanding of system performance.

System Calibration and Improvement: By regularly assessing MAPE, researchers and engineers can identify areas where the eye tracking system may be less accurate. This insight is invaluable for calibration and continual improvement of the system.

User Experience Optimization: In applications like web design and marketing, understanding where users are looking is crucial. A lower MAPE ensures that designers can rely on eye tracking data for optimizing user interfaces and content.

Research Validity: In scientific research, especially in fields like psychology and neuroscience, the validity of experimental results often hinges on the accuracy of eye tracking. A low MAPE is indicative of a more reliable system, enhancing the validity of research outcomes.

5.4 Challenges and Considerations

While MAPE is a powerful metric, it's essential to consider certain challenges and nuances:

Data Quality: MAPE is only as reliable as the quality of the data it is based on. Poor data quality can lead to inaccurate assessments.

Participant Variability: Individuals may have unique eye movement patterns, making it challenging to achieve low MAPE values for all users consistently.

Task Dependency: MAPE may vary based on the complexity of the task. Certain tasks may inherently have higher prediction errors.

Environmental Factors: Lighting conditions, screen calibration, and participant positioning can influence MAPE. These factors need careful consideration during data collection.

6. Accuracy Results

6.1 Claimed Accuracy

Our stated accuracy stands at **0.7** MAPE, assessed based on the criteria outlined in the table below.

MAPE	Forecasting power
<10%	Highly accurate forecasting
10% - 20%	Good forecasting
20% - 50%	Reasonable forecasting
>50%	Weak and inaccurate forecasting

Table 6.1 User test flow

7. Evaluation Metrics for Eye Tracking Devices and Algorithms

Assessing the accuracy and performance of eye tracking devices and algorithms is vital for understanding the capabilities and limitations of the system. These metrics allow us to delve into the intricate characteristics of gaze tracking systems and their responses to variable operating conditions.

7.1 Angular accuracy metrics

Processing extracts the respondent's gaze position expressed in the camera coordinate system, relying on deep learning models to estimate the gaze position from the webcam video. The models used in this processing step are pre-trained on thousands of faces using deep learning algorithms.

7.1.1 Gaze angular accuracy

Gaze Angular Accuracy refers to the precision and correctness of a gaze tracking system in determining the direction of a person's gaze in relation to a reference point on a screen. It is measured in degrees and indicates how closely the system's estimated gaze direction aligns with the actual gaze direction of the user.

It helps to determine how accurately a gaze tracking system can determine where a person is looking. The lower the angular accuracy value, the more precise and accurate the system is in determining the user's gaze direction. High angular accuracy is crucial for applications such as virtual reality, human-computer interaction, and usability studies where precise gaze information is essential.

To ensure consistency and comparability in evaluating eye tracking systems, a common set of accuracy metrics is essential. Gaze angular accuracy propose a standardized approach to measure and compare eye tracker accuracy using raw data and ground truth locations. Experiments reveal that the mean angular accuracy varies with user-tracker distance, demonstrating the critical role this factor plays in gaze accuracy.

**Source: Eye tracking methodology: Theory and practice. Springer Science & Business Media.

7.1.2 Gaze Yaw and Pitch Angular Accuracies

Gaze Yaw Angular Accuracy refers to the horizontal movement of the gaze, typically measured in degrees. Gaze yaw angular accuracy indicates how accurately a gaze-tracking system can determine the left or right direction of a person's gaze. For example, if the system indicates that the user is looking 30 degrees to the right, the accuracy is determined by how closely this aligns with the actual gaze direction.

Gaze Pitch Angular Accuracy, on the other hand, represents the vertical movement of the gaze. Gaze pitch angular accuracy assesses the precision of the system in determining whether a person is looking up or down. It is also measured in degrees and reflects how closely the system's estimation matches the actual pitch angle of the user's gaze.

Fig 7.1 (a)Gaze Yaw and Pitch Angular Accuracies (b) Shows the different orientations of the eye tracker in neutral position (top left) and (clockwise from top) roll, pitch and yaw variations. (c) Shows the gaze error sensitivity to orientations of the tablet when eye tracking is performed on it.

7.2 Statistical accuracy metrics

7.2.1 Statistical Measures of Eye Tracking Performance

Statistical measures of eye tracking performance refer to quantitative metrics used to assess the accuracy, reliability, and effectiveness of eye tracking systems. These measures are essential in evaluating how well such systems capture and interpret users' eye movements.

7.2.2 Histogram-Based Metrics

Histogram-based metrics offer a means to compare data from different eye trackers or experiments. Metrics like correlation, intersection, and Bhattacharya distance aid in quantifying data similarity. By applying these metrics, we can assess the similarity between gaze error data from different experiments, shedding light on the performance of eye tracking systems.

7.3 Sensitivity metrics

7.3.1 Head Pose Sensitivity

Variations in a user's head pose can impact gaze tracking accuracy, yet quantifying this effect is challenging. Our experiments provide valuable insights into how head pose influences the accuracy of the tracker. For example, results reveal that head pose sensitivity analysis significantly affects accuracy, highlighting the need for head constraint devices in certain cases.

7.3.2 Platform Orientation Sensitivity

The impact of platform orientation on mobile devices such as tablets and smartphones is often overlooked in eye tracking studies. We delve into this uncharted territory and offer insights into how platform orientation affects tracking accuracy. Our analysis shows that pitch angle variations have the most substantial impact on gaze errors when using tablet-based eye trackers.

**Source: https://www.researchgate.net/

7.3.3 Gaze Tracking Efficiency

Metrics for gaze tracking efficiency help identify the optimal user distance and viewing angles for a given tracker. These metrics aid in understanding where the best tracking accuracy can be achieved. Our findings suggest that the ideal user distance for reliable gaze tracking falls within a specific range, which varies depending on the tracker used.

7.3.4 Error Spatial Density

Error magnitudes, alone, do not provide information about the spatial distribution of gaze errors. To address this, we introduce a spatial error density metric to visualize gaze error distribution on the screen. This metric is crucial for understanding if specific screen regions are more prone to errors, enabling targeted improvements in tracking accuracy.

7.3.5 Gaze Error Analysis with Respect to Visual Eccentricity

The location of gaze targets on the display screen can significantly affect tracking quality. We evaluate gaze errors concerning stimulus eccentricity, providing insights into tracker performance under various conditions. Our analysis indicates that central regions of low visual angles yield more reliable results, while screen corners exhibit higher error levels for most users.

7.3.6 The ROC Metric: Subjective Performance Evaluation of Eye Tracking Systems

We introduce the concept of a subjective performance evaluation metric to supplement traditional objective metrics. This metric allows for the evaluation of tracker performance at varying accuracy thresholds. Our ROC analysis demonstrates how the performance of an eye tracker can change when different accuracy thresholds are applied, offering a more nuanced view of its capabilities.

8. Visualization of Eye Tracking Metrics

8.1 Heatmaps

Heat maps are static or dynamic aggregations of gaze points revealing the distribution of visual attention. Following an easy-to-read color-coded scheme - with red areas suggesting a high number of gaze points (and therefore an increased level of interest), and yellow and green areas showing fewer gaze points (a less engaged visual system). Areas without coloring were likely not attended to at all.

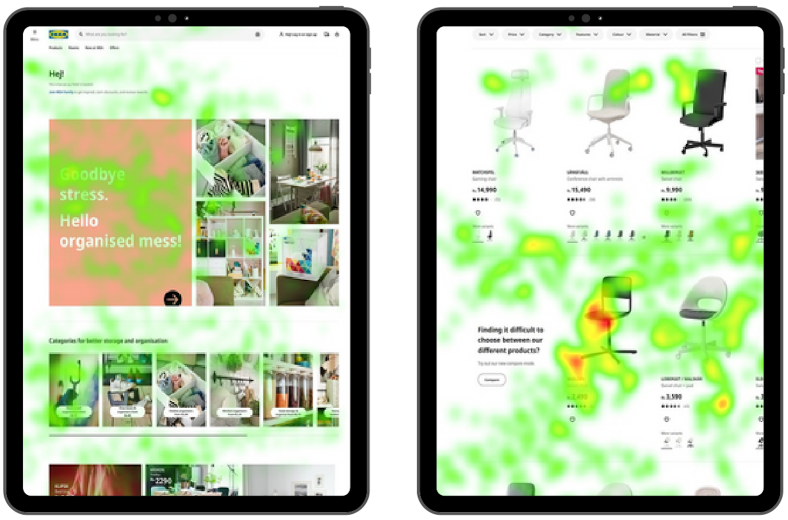


Fig 8.1 Heatmaps

**Source: Eye tracking methodology: Theory and practice. Springer Science & Business Media.

8.2 Time to First Fixation (TTFF)

The time to the first fixation indicates the amount of time it takes a respondent to look at a specific AOI from stimulus onset. TTFF can indicate both bottom-up stimulus-driven searches (Eg. A flashy company label catching immediate attention) as well as top-down attention-driven searches (Eg. respondents actively decide to search for certain elements or areas on a website).

8.3 Fixation Sequences

Based on fixation position (where?) and timing information (when?) you can generate a fixation sequence. This is dependent on where respondents look and how much time they spend, and provides insight into the order of attention, telling you where respondents looked first, second, third and so on

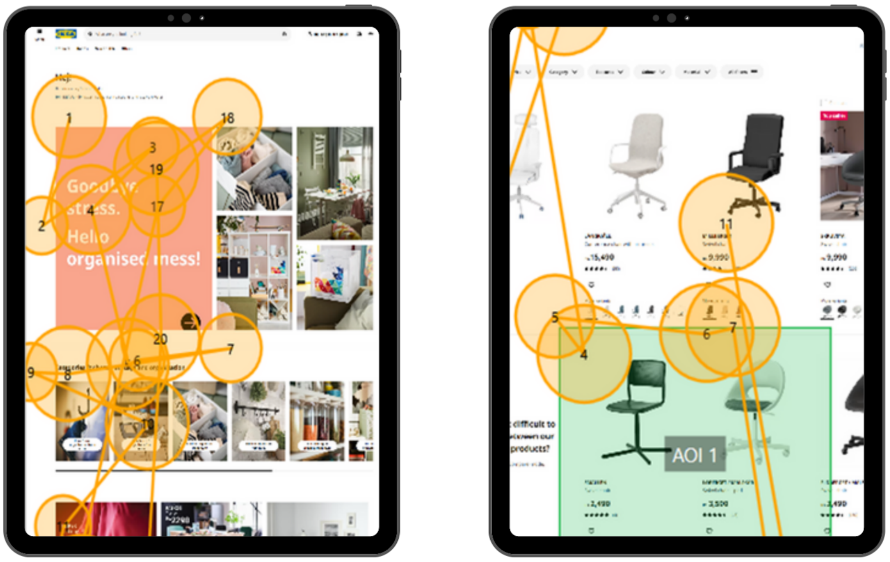


Fig 8.2 Fixation maps

8.4 Areas Of Interests (AOIs)

Areas of Interest also referred to as AOIs, are user-defined subregions of a displayed stimulus. Extracting metrics for separate AOIs might come in handy when evaluating the performance of two or more specific areas in the same video, picture, website, or interface. Eg: compare groups of participants, conditions, or different features within the same scene.

AOI metrics	AOI 1	AOI 2
Information		
AOI duration (ms)	6007.2	6007.2
AOI duration (%)	100	100
Size (cm2)	233.7	40.9
Size (%)	11.6	2
Respondent base	15	15
Fixation based metrics		
Respondent ratio (%)	93.3	26.7
Revisit count	1.1	0
Fixation count	4.5	1.3
TTFF AOI (ms)	690.3	3523.1
Dwell time (ms)	2061.8	707
Dwell time (%)	34.3	11.8
First fixation duration (ms)	514.4	638.9
Saccade based metrics		
Respondent ratio (%)	86.7	6.7
Saccade count	2.8	1
Amplitude (deg)	5.8	2
Peak velocity (deg/s)	113.6	60.9

Fig 8.3 Area of Interest (AOIs)

8.5 Time Spent

Time spent quantifies the amount of time that respondents spent looking at an AOI. As respondents have to blend out other stimuli in their visual range that could be equally interesting, the amount of time spent often indicates motivation and conscious attention (prolonged visual attention at a certain region clearly points to a high level of interest).

8.6 Respondent Count

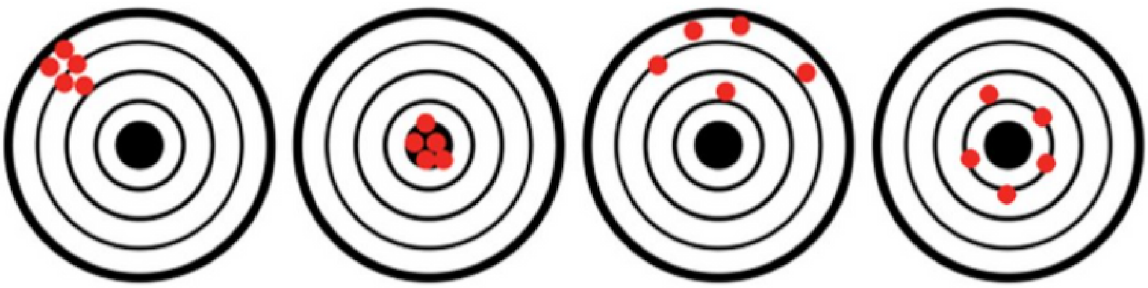
The respondent count describes how many of your respondents actually guided their gaze towards a specific AOI. A higher count shows that the stimulus is widely attended to, while a low count shows that little attention is paid to it.

9. Understanding Accuracy and Errors

9.1 How Accurate if Eye Tracking

The best-case scenario is high precision and accuracy. A system with good accuracy and precision will provide more valid data because it can accurately describe the location of a person's gaze. However, any combination of precision and accuracy is possible. Accuracy and precision are measured in degrees of the visual angle after the calibration. The lower the value, the better its accuracy and precision.

The level of accuracy and precision required is determined by the nature of the eye-tracking study. Small uncertainties, for instance, can be critical when analyzing gaze data in reading studies or studies with a small stimulus. Real-world performance depends on many factors, such as room lighting, participant eyewear, and testing space setup. Careful consideration and the application of best practices can help in obtaining the highest quality eye-tracking data possible.



Poor accuracy but good precision | **Good accuracy and good precision** | **Poor accuracy and poor precision** | **Good accuracy but poor precision**

Fig 9.1 Depiction of accuracy with precision in eye tracking

10. Best Practices for Improving Accuracy

Failure or complications in studies most often occur due to small mistakes that could have easily been avoided. Often this happens because participants just didn‘t know about the basics to avoid running into issues.

Environment and Lighting Conditions

For eye tracking, lighting conditions are essential. Avoid direct sunlight coming through windows as sunlight contains infrared light that can affect the quality of the eye-tracking measurements. Avoid brightly lit rooms (overhead light). Ideally, use ambient light.

**Source: Imotions Eye Tracking Guide.

Respondent should stay within the "headbox"

If the respondent moves too far away from the eye tracker, the camera will not be able to reliably detect the eyes. Make sure that the respondent is sitting comfortably in front of the eye tracker, and that they are aware that they shouldn't move around too much.

Avoid Obstructions

Eye tracking requires a clean line of sight between the eyes and the camera. This seems obvious for an entire experiment, but even brief interruptions can take away crucial data.
While the first step in preventing interruptions is to make sure that respondent knows not to obstruct their eyes, the specifications of the eye tracker can also have an impact.

Good calibration is essential

If the device calibrates to the wrong measurements, it will take those errors into the results which could render your experiment meaningless. This is obviously a situation you want to avoid.

Data Quality

Precise, accurate data ensures reliable insights into user behavior. Rigorous validation processes are key for meaningful and impactful eye tracking analyses.

Ensure all people involved are properly trained

It is essential that the people involved in data collection are trained on the systems used so that they have a level of knowledge that allows them to run a study smoothly. Having to train people before the testing process is advantageous and allows you to prepare for potential mishaps or missteps made during the experiment.

10.1 Eye Tracking: Pros

- **Data Collection Scalability:** Anyone with a functioning webcam can potentially join your study from anywhere in the world, making quantitative data collection much easier.
- **Price:** With thelightBulb.ai's core software and Online Data Collection modules you can conduct studies using only a computer capable of running the software. No additional hardware is required – aside from your respondents' webcams.

11. Conclusion

The data quality of webcam eye-tracking is not comparable to that of dedicated eye-tracking hardware, it is an excellent choice for scaling your research. We like to think of it as giving our clients the ability to conduct quantitative human behavior research for the first time.

If you plan to conduct bulk UX testing, A/B testing, or image/video studies with eye tracking, we are confident that our eye-tracking feature will be a useful tool for you.

Thelightbulb.ai uses latest in computer vision, emotion ai & machine learning technologies to measure attention and emotions of opt-in participants as they consume content & experiences online.

LET'S TALK!

SALES@THELIGHTBULB.AI