Preregistration for Quantitative Research in Psychology (PRP-QUANT) Template

Title

T1 Title
The title should be focused and descriptive, using relevant key terms to reflect what will be done in the study. Use title case (https://apastyle.apa.org/style-grammar-guidelines/capitalization/title-case).
Using a test battery to compare three remote, video-based eye-trackers

T2 Contributors, Affiliations, and Persistent IDs (recommend ORCID iD)
Provide in separate entries the full name of each contributor, each contributor's professional affiliation, and each contributor's persistent ID. See ORCID iD for an example of persistent ID (https://orcid.org/). Optional: include the intended contribution of each person listed (e.g. statistical analysis, data collection; see CRediT, https://casrai.org/credit/).
Lisa Spitzer[1], Stefanie Mueller[1] [1]Leibniz Institute for Psychology (ZPID)

T3 Date of Preregistration
This is assigned by the system upon preregistration submission.
17/09/2021

T4 Versioning information

This is assigned by the system upon submission of original and subsequent revisions. Should be a persistent identifier, if not a DOI.

T5 Identifier

This unique identifier is assigned by the system upon submission.

T6 Estimated duration of project

Include best estimate for how long the project will take from preregistration submission to project completion.

Due to the current Covid-19 pandemic, the time it will take to collect the data is rather unpredictable. However, it is planned to collect and analyze all data within a year.

T7 IRB Status (Institutional Review Board/Independent Ethics Committee/Ethical Review Board/Research Ethics Board)

If the study will include human or animal subjects, provide a brief overview of plans for the treatment of those subjects in accordance with established ethical guidelines. If appropriate institutional approval has been obtained for the study, provide the relevant identifier here. If the study will be exempt from ethical board review, provide reasoning here.

The study was approved by the ethics committee of Trier University, Germany.

T8 Conflict of Interest Statement

Identify any real or perceived conflicts of interest with this study execution. For example, any interests or activities that might be seen as influencing the research (e.g., financial interests in a test or procedure, funding by pharmaceutical companies for research).

The authors declare that they have no conflict of interest.

T9 Keywords

Include terms specific to your topic, methodology, and population. Use natural language and avoid words used in the title or overly general terms. If you need help with keywords, try a keyword search using your proposed keywords in a search engine to check results.

reproducibility, eye-tracking, Tobii, Gazepoint, EyeLink

T10 Data accessibility statement and planned repository

"We plan to make the data available (yes / no)
If "yes", please specify the planned data availability level by selecting one of the options:

- Data access via download; usage of data for all purposes (public use file)
- Data access via download; usage of data restricted to scientific purposes (scientific use file)
- Data access via download; usage of data has to be agreed and defined on an individual case basis
- Data access via secure data center (no download, usage/analysis only in a secure data center)
- Data available upon email request by member of scientific community
- Other (please specify)

We plan to make the data available in PsychArchives. Additionally, a screen recording of the experimental procedure will be shared publicly.
Data access via download; usage of data for all purposes (public use file)

T11 Optional: Code availability

We plan to make the code available (yes / no).
If "yes", please specify the planned code availability level (use same descriptors of data in T10).

We plan to make the code available.
Data access via download; usage of data for all purposes (public use file).

T12 Optional: Standard lab practices

Standard lab practices refer to a (timestamped) document, software package, or similar, which specifies standard pipelines, analytical decisions, etc. which always apply to certain types of research in a lab. Specify here and refer to at the appropriate positions in the remainder of the template:
We plan to make the standard lab practices available (yes / no).
If "yes", please specify the planned standard lab practices availability level (use same descriptors of data in T10).

No, we do not have a document specifying standard lab practices for this kind of studies.

Abstract

(150 words)

A1 Background

(See introduction I1)

Eye-tracking research is based on data collected with a multitude of eye-tracking devices. The comparability of results obtained by different devices is unclear.

A2 Objectives and Research questions

(See introduction I2)

This study aims to investigate the comparability of data of three remote eye-tracking devices, the Gazepoint GP3HD Desktop, the Tobii Pro X3-120, and the EyeLink 1000+, using an extensive test battery.

A3 Participants

(See methods M4)

Data of *N* = 25 participants will be collected, who will be recruited by subject pool advertisement, email lists, and personal invitation. Inclusion criteria comprise a successful calibration with all three eye-tracking devices. Participants will receive monetary compensation.

A4 Study method

(See methods M10-14)

Participants will complete eight eye-movement tasks, consecutively measured by each eye-tracker. The tasks measure various eye-tracking parameters, such as accuracy, precision, smooth pursuit, microsaccades, or pupil dilation. The measure of data quality obtained in each task, will be compared between devices.

Introduction

(no word limit)

I1 Theoretical background

Provide a brief overview that justifies the research hypotheses.

- A large number of eye-trackers, developed by various manufacturers, are available, for example:
  - Tobii (https://www.tobii.com/)
  - SR Research (https://www.sr-research.com/)
  - Gazepoint (https://www.gazept.com/)
  - *For a list of devices also see:* https://www.lboro.ac.uk/microsites/research/applied-vision/projects/vision_resources/emed.htm
- There exist a number of empirical studies in which different devices are compared with each other, to assess how comparable the obtained data quality and the results are:
  - Dalmaijer (2014) compared a low-cost and a high-cost eye-tracker (EyeTribe vs. EyeLink 1000); the results indicated that the spatial precision and accuracy of the lower-cost eye-tracker are good enough for some measures, but unsuitable for testing high-accuracy saccade metrics
  - Holmqvist (2017) compared 12 remote and tower-mounted eye-trackers, and reported deviating values for accuracy and precision between the eye-trackers, as well as various influences on the data quality (e.g., person characteristics)
  - Macinnes et al. (2018) compared three wearable eye-trackers (i.e., Pupil Labs 120Hz Binocular glasses, SMI ETG 2 glasses, and the Tobii Pro Glasses 2); in this study, the Pupil Labs showed better accuracy than both the SMI and Tobii eye-trackers; and the Tobii eye-tracker was significantly less precise than the Pupil Labs or SMI eye-trackers
- The present study extends these results by comparing two eye trackers that have not yet been part of detailed comparisons with other devices. The Eye Link 1000+ serves as a reference. Tobii, Gazepoint and SR Research are among the most well-known eye-tracking manufacturers (see https://imotions.com/blog/top-eye-tracking-hardware-companies/)
  - *Tobii*: Tobii is the current market leader in eye-tracking devices. The model to be examined here is the Tobii Pro X3-120, which can record at a sampling rate of 120 Hz with an accuracy of 0.4° and precision of 0.24° when measured binocularly at ideal conditions.

- *Gazepoint*: The next considered eye-tracker is the Gazepoint GP3HD Desktop. This model measures at 150 Hz and is therefore comparable to the Tobii, but is considerably less expensive. The reported accuracy is between 0.5 – 1°. This eye-tracker has been evaluated before in the context of psychophysiology (Cuve et al., 2021), however, a direct comparison with other eye-trackers to directly assess the reproducibility of results and comparability of data quality is missing
- EyeLink: SR Research's EyeLink is considered the gold standard in eye-tracking research (https://www.sr-research.com/) because it is the most accurate (0.25º – 0.50º typical) and precise (0.01º RMS if head is supported) eye-tracker and has an extremely high sampling rate of up to 2000 Hz.
- These devices are also available in PsychLab Offline, an open science service provided by the Leibniz Institute for Psychology (ZPID, https://leibniz-psychology.org/en/services/data-collection/psychlab-offline/):
    - Researchers can apply with their study to receive free data collection based on their proposed study idea
    - Authors can choose between these three devices in the lab
    - To allow a more informed decision, it is beneficial to provide detailed information about which eye-tracker is particularly suitable for which specific study situation or for which measured parameter

- Previous studies comparing different devices mainly focused on accuracy and precision of gaze position when assessing the quality of eye tracking data
- However, accuracy and precision alone are not sufficient to serve as a benchmark for an eye-tracker, since a wide range of measures can be obtained with eye-tracking, for example:
    - Classification of events as blinks, saccades, glissades, fixations (which is possible by using various different algorithms, for example the one by Nyström & Holmqvist, 2010)
    - Smooth pursuit movements (e.g., see Robinson, 1965)
    - Microsaccades (e.g., see Engbert & Mergenthaler, 2006)
    - Pupil dilation, which is, for example, associated with cognitive effort (e.g., see van der Wel & van Steenbergen, 2018)
    - Influences of different positions on data quality (e.g., see Niehorster et al., 2018)
- The present study will assess these parameters by using an extensive test battery developed by Ehinger et al (2019):
    - Measures "fixation and saccade properties in an artificial grid and in a free-viewing task, decay of accuracy, smooth pursuit, pupil dilation, microsaccades, blink detection, and the influence of head motion" (Ehinger et al., p. 2)
    - Was tested on the Pupil Labs glasses and the EyeLink 1000

- This test battery has been adapted for the current study to compare three remote, video-based eye-trackers: Gazepoint GP3HD Desktop, Tobii Pro X3-120, and EyeLink 1000+

I2 Objectives and Research question(s)

Outline objectives and research questions that inform the methodology and analyses (below).

Objective:
Quantitative comparison for a wide range of data quality measures between the eye-trackers Gazepoint GP3HD Desktop, Tobii Pro X3-120, and EyeLink 1000+.

Research questions:
To what extent does the eye-tracking device used to measure a parameter influence the measurement outcome, that is, are there significant differences and how similar are results?
Specifically, we will investigate the following parameters:

- Accuracy, precision, and their decay over time under standard conditions
- Accuracy and precision under suboptimal conditions (i.e., influence of pupil dilation and head position)
- Point of gaze measurement
- Smooth pursuit
- Microsaccades
- Pupil dilation

I3 Hypothesis (H1, H2, …)

Provide hypothesis for predicted results. If multiple hypotheses, uniquely number them (e.g., H1, H2a, H2b,) and refer to them the same way at other points in the registration document and in the manuscript.

Rather than being based on a theory from which testable hypotheses can be derived, the present study focuses on technical aspects and their potential methodological consequences for eye-tracking research. Thus, rather than hypotheses with certain predictions, we plan conducting a number of comparisons to determine differences between measurements obtained by different devices. The following parameters will be assessed and compared between the three eye-trackers:

1. **Accuracy and precision**
   - 1.1. Overall accuracy and accuracy decay (task 1, 4, 6)
   - 1.2. Overall precision and precision decay (task 1, 4, 6)
   - 1.3. Accuracy and precision under differing luminance conditions (task 5)
   - 1.4. Accuracy and precision under suboptimal (position) conditions (task 7, 8)

2. Point of gaze measurement (X and Y coordinates)

2.1. Mean fixation locations under standard conditions (task 1)
2.2. Mean fixation locations with varying pupil dilations (task 5)
2.3. Mean fixation locations under suboptimal position conditions (task 7, 8)
2.4. Horizontal and vertical bias under suboptimal position conditions (task 7, 8)

3. Other important eye-tracking parameters

3.1. Smooth pursuit movements (task 2)
3.2. Microsaccades (task 3)
3.3. Pupil dilation (task 5)

I4 Exploratory research questions (if applicable; E1, E2, ...)

If planning exploratory analyses, provide rationale for them here. If multiple exploratory analyses, uniquely number them (E1, E2, ...) and refer to them in the same way in the registration document and in future publications.

Method

M1 Time point of registration
Select one of the options: ● Registration prior to creation of data ● Registration prior to any human observation of the data ● Registration prior to accessing the data ● Registration prior to analysis of the data ● Other (please specify; might include if T1 longitudinal data has been analyzed, but T2 has not yet been analyzed)
Registration prior to creation of data

M2 Proposal: Use of pre-existing data (re-analysis or secondary data analysis)
Will pre-existing data be used in the planned study? If yes, indicate if the data were previously published and specify the source of the data (e.g., DOI or APA style reference of original publication). Specify your level of knowledge of the data (e.g., descriptive statistics from previous publications), whether or not this is relevant for the hypotheses of the present study, and how it is assured that you are unaware of results or statistical patterns in the data of relevance to the present hypotheses.
No pre-existing data will be used.

Sampling Procedure and Data Collection

M3 Sample size, power and precision
(1) Relevant sample sizes: e.g., single groups, multiple groups, and sample sizes (or sample ranges) found at each level of multilevel data. (2) Provide power analysis (e.g. power curves) for fixed-N designs. For sequential designs, indicate your 'stopping rule' such as the points at which you intend to be viewing your data and in any way analyzing them (e.g., t-tests and correlations, but even descriptively such as with histograms).
Mixed models are used to compare the different eye-trackers in a repeated measures design. As a primer for power, a power analysis was conducted for one of our main analyses: Analysis 1.1 model A (accuracy). In order for G*Power (Faul, Erdfelder, Lang & Buchner, 2007; Faul, Erdfelder, Buchner & Lang, 2009) to be used for the power analysis, a very related method to the Mixed Model was used, i.e., repeated measures ANOVA.

Our pilot data (*N* = 4) showed very high effect sizes for the main effect of eye-tracker in the mixed model of Analysis 1.1 model A (i.e., dependent variable: accuracy; fixed effects: eye-tracker, target, eye-tracker * target; random effect: participant), η^2 = .276. However, since the small pilot data set of only four subjects, some of whom showed very high accuracy values in the later tasks, cannot be assumed to be representative, a smaller effect size was used for the power analysis, namely η^2 = .06 , in order to be able to detect medium sized effects (Cohen, 1988).

F tests - ANOVA: Repeated measures, within factors (option - “as in SPSS”)
Analysis: A priori: Compute required sample size
Input: Effect size f(U) = 0.2526456
α err prob = 0.05
Power (1-β err prob) = 0.95
Number of groups = 1
Number of measurements = 147
Nonsphericity correction ε = 0.75
Output: Noncentrality parameter λ = 69.8936301
Critical F = 1.2473982
Numerator df = 109.5
Denominator df = 1095
Total sample size = 11
Actual power = 0.9745941

Thus, in order to be able to detect medium sized effects for the main effect of eye-tracker in the accuracy/accuracy decay mixed model (with α = 0.05 and with a power of 95%), the targeted sample size should be *N* = 11. However, to be sure not to miss any effects and since this is more in line with other eye-tracking studies, we will collect *N* = 25 subjects.

M4 Participant recruitment, selection, and compensation

Indicate (a) methods of recruitment (e.g., subject pool advertisement, community events, crowdsourcing platforms, snowball sampling); (b) selection and inclusion/exclusion criteria (e.g., age, visual acuity, language facility); (c) details of any stratification sampling used; (d) planned participant characteristics (gender, race/ethnicity, sexual orientation and gender identity, SES, education level, age, disability or health status, geographic location); (e) compensation amount and method (e.g., same payment to all, pay based on performance, lottery).

A) Method of recruitment: Subject pool advertisement, email lists, and personal invitation
B) Selection and inclusion/exclusion criteria:
   - No glasses or hard contact lenses (people with glasses can participate if they can wear soft contact lenses for the time of the experiment)
   - Calibration accuracy limits (validation accuracy = 2°) need to be met at the beginning of the study and before every block for a participant to be included
C) Stratification sampling: no

D) Planned participant characteristics: Based on the planned recruitment, it can be assumed that mainly data of healthy young adults with high education level (e.g., students of the University of Trier) will be collected.
E) Compensation: Participants will receive a monetary compensation of 10 € per hour, thus approximately 15 € for the duration of the experiment.

M5 How will participant drop-out be handled?

Indicate any special treatment for participants who drop out (e.g., there is follow-up in a manner different from the main sample, last value carried forward) or whether participants are replaced.

Participants will be replaced.

M6 Masking of participants and researchers

Indicate all forms of masking and/or allocation concealment (e.g., administrators, data collectors, raters, confederates are unaware of the condition to which participants were assigned).

No masking will be used. Participants will be explained the study's purpose before beginning the study. The participant and the administrator will know the current eye-tracker condition based on the set-up. Furthermore, the analyst will know which data has been collected with each eye-tracker as this information is included in the data.

M7 Data cleaning and screening

Indicate all steps related to data quality control, e.g., outlier treatment, identification of missing data, checks for normality, etc.

Eye-trackers will be calibrated at the beginning of the experiment and before each block. After the calibration, a manually programmed validation procedure will be used to assess the calibration quality. Here, a cut-off accuracy of 2° will be used. Thus, participants will only be measured if calibration is successful and their validation accuracy is $\leqq 2°$. In case of problems with the calibration, three attempts will be made for each block to complete the calibration successfully. If the calibration still does not succeed, the data collection for this participant will be aborted and they will be replaced.

M8 How will missing data be handled?

Indicate any procedures that will be applied during the analysis to deal with missing data, such as (a) case deletions; (b) averaging across scale items (to handle missing items for some); (c) test of missingness (MAR, MCAR, MNAR assumptions; (d) imputation procedures (FIML vs. MI); (e) Intention to treat analysis and per protocol analysis (as appropriate).

If more than >25% of data within a trial is corrupt (i.e., point of gaze validity is not given), this trial will be excluded post data collection (see AP2).
NA measurements will be used during analysis to identify blinks. Remaining corrupt data 1) which indicates gaze outside the monitor frame, 2) data identified as “invalid” by the eye-tracking device, e.g., NA time stamps, and 3) data for which velocities are >1000 degrees/s or >100.000 degrees/s² (as recommended by Nyström & Holmqvist, 2010) will be excluded from the data (see AP3).

M9 Other information (optional)

For example, training of raters/participants or anything else not yet specified.

Conditions and design

M10 Type of study and study design

Indicate the type of study (e.g., experimental, observational, crosssectional vs. longitudinal, single case, clinical trial) and planned study design (e.g., between vs. within subjects, factorial, repeated measures, etc.), number of factors and factor levels, etc..

Type of study: experimental
Planned study design: within-subjects design / repeated measures: every participant is tested with each of the three eye-trackers and completes each of the eight tasks listed below.
The tasks are based on the study by Ehinger et al. (2019) and were adapted for this study:

Task 1 - Grid 1:
Participants are shown 49 fixation targets which are presented consecutively at different locations on the screen, and are asked to fixate each target and to press the spacebar to confirm their fixation (after each confirmation, the next fixation target is shown).

Task 2 - Smooth pursuit:
In eight trials, bullseye targets are presented which move across the screen (starting from the center). At the beginning of each trial, participants see a target in the center of the screen that they are asked to fixate. They press the spacebar to start the movement. After a short delay, the bullseye will move across the screen, and participants are asked to follow it with their eyes as long as possible.

Task 3 - Microsaccades:
A bullseye target is shown in the center of the screen. Participants are asked to fixate it until it disappears after 20s.

Task 4 - Grid 2:
This task is similar to task 1.

Task 5 - Pupil dilation grid:
Participants are shown bullseye targets at different locations on the screen while the luminance of the background is varied in each trial. The five targets are a subset of targets shown in the large grid tasks (task 1, 4, 6). Each of the targets is shown with each of the four luminance conditions (i.e., 20 trials in total). Between trials, a black baseline background is shown.

Task 6 - Grid 3:
This task is similar to task 1.

Task 7 - Tilt position:
A line along with a fixation target are presented in the center of the screen. Participants are asked to tilt their head so that their eyes align with the presented line while they still fixate the target, and to press the spacebar to confirm the trial (seven trials in total).

Task 8 - Turn position:
A fixation target is presented at different locations on the screen. Participants are asked to turn their head so that their nose points to the target and to fixate it. Once they press the spacebar to confirm the position, the next target will appear (seven trials in total). The seven targets of this task are a subset of targets shown in the large grid tasks (task 1, 4, 6).

M11 Randomization of participants and/or experimental materials

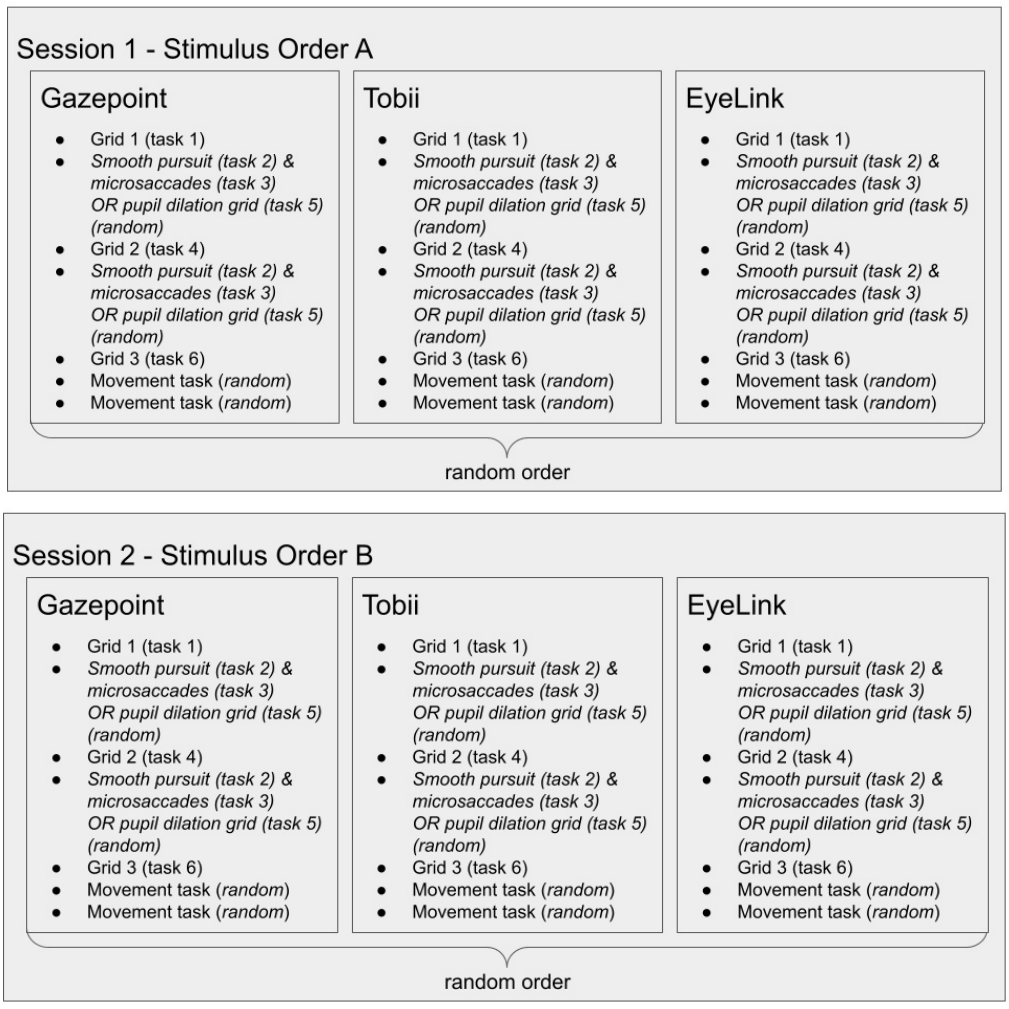
If applicable, describe how participants are assigned to conditions or treatments, how stimuli are assigned to conditions, and how presentation of tests, trials, etc. is randomized. Indicate the randomization technique and whether constraints were applied (pseudo-randomization). Indicate any type of balancing across participants (e.g., assignments of responses to hands, etc.).

Pseudo-randomization of task order:
All participants complete two blocks per eye-tracking device (six in total). Each block comprises eight tasks (see M10). These tasks are shown in a pseudo-randomized order, i.e., some of the tasks are presented at a fixed time point, whereas others are presented at a random time point. Specifically, the grid tasks (task 1, 4, and 6) are always presented at the beginning (task 1), in the middle (task 4) and at the end (task 6) of each block because they are used to measure accuracy decay. Furthermore, the two position tasks (task 7 and 8) are always presented after all other tasks in a block to prevent the calibration in the other tasks from being affected by the head movement. Since the pupil dilation grid task (task 5) is estimated to be the longest task, it is shown alone between two grid tasks (i.e., between task 1-4 or 4-6), while task 2 (smooth pursuit) and 3 (microsaccades) are presented together between the respective other two grid tasks. Which of these (task 2 and 3 vs. task 6) is presented first is decided randomly (see Figure 1).

(Pseudo-)Randomization of stimuli order within a task:
In the grid tasks, the targets are shown in a random order. Meanwhile, for the other tasks the stimuli order within the tasks has been pseudo-randomized according to specific criteria. Possible stimulus characteristics (e.g., coordinates, luminances of the screen, delay durations) were shuffled at the planning stage of the study to create specific stimuli orders. For task 5 (pupil dilation grid), the order was altered manually after shuffling to ensure that two identical target positions / luminances are not shown directly after each other. Overall, two different pseudo-randomized sets of stimuli orders for all tasks were created (A, B). Each of these resembles one block with a specific stimuli order for all pseudo-randomized tasks (see Figure 1).

Sequential presentation of the stimuli orders, randomized order of the eye-trackers:
Both sets (A, B) are measured in turn with each of the three eye-trackers to ensure high comparability. To avoid fatigue, the data collection is divided into two sessions. In each session, one stimuli order is measured in sequential order for each eye tracker. That is, in the first session, stimuli order A is measured with the three eye-trackers one after the other (resulting in the measurement of three blocks in total). In session 2, stimuli order B is measured with the three eye-trackers in succession (i.e., again, three blocks are measured). The order in which the eye-trackers are measured is decided upon the beginning of each session randomly (see Figure 1).

Figure 1
Overview over (pseudo-)randomized order of sessions, stimuli orders, eye-trackers, and tasks



M12 Measured variables, manipulated variables, covariates

This section shall be used to unambiguously clarify which variables are used to operationalize the hypotheses specified above (item I3). Please (a) list all measured variables, and (b) explicitly state the functional role of each variable (i.e., independent variable, dependent variable, covariate, mediator, moderator). It is important to (c) specify for each hypothesis how it is operationalized, i.e., which variables will be used to test the respective hypothesis and how the hypothesis will be operationally defined in terms of these variables. The description here shall be consistent with the statistical analysis plans specified under AP6 (below).

Eye-tracking data of three eye-tracking devices will be recorded. The output format of the raw data depends on built-in software of the eye tracker when used with our experimental program (OpenSesame). This is different for each eye-tracker. We will transform the recorded data of each eye-tracking device into a standard format including the following variables:

timestamp	Timestamp in seconds
pog_x_left	X coordinate of left eye in cm (0 = center of screen, negative values = left, positive values = right)
pog_y_left	Y coordinate of left eye in cm (0 = center of screen, negative values = up, positive values = down)
pog_x_right	X coordinate of right eye in cm (0 = center of screen, negative values = left, positive values = right)
pog_y_right	Y coordinate of right eye in cm (0 = center of screen, negative values = up, positive values = down)
pog_x	X coordinate of both eyes in pixels (mean if data of both eyes are available, otherwise this is the valid datapoint of either left or right eye)
pog_y	Y coordinate of both eyes in pixels (mean if data of both eyes are available, otherwise this is the valid datapoint of either left or right eye)

pupilsize_left	Pupil size of left eye as a fraction of size in comparison to maximum value
pupilsize_right	Pupil size of right eye as a fraction of size in comparison to maximum value
pog_validity_left	Validity of point of gaze data of left eye (1 = valid, 0 = not valid)
pog_validity_right	Validity of point of gaze data of right eye (1 = valid, 0 = not valid)
pog_validity	Validity of point of gaze data of both eyes (1 = valid, 0 = not valid) → is valid if at least one of left or right POG is valid
pupilsize_validity_left	Validity of pupil size of left eye (1 = valid, 0 = not valid)
pupilsize_validity_right	Validity of pupil size of right eye (1 = valid, 0 = not valid)
task	Task 1 = Grid 1 2 = Smooth pursuit 3 = Microsaccades 4 = Grid 2 5 = Pupil dilation grid 6 = Grid 3 7 = Tilt position 8 = Turn position
trial	Trial For task 1: 49 trials For task 2: 8 trials For task 3: 1 trial For task 4: 49 trials For task 5: 20 trials For task 6: 49 trials For task 7: 7 trials For task 8: 7 trials

Furthermore, information about the experimental context will be saved:

- trial start and stop
- position of target
- direction, speed, and delay of smooth pursuit target (task 2)
- duration of black screens and luminance screens (task 5)
- induced head rotation (task 7)

These variables will serve as basis for the calculation of further parameters (e.g., for the identification of blinks, saccades, glissades, and fixations).

M13 Study Materials

Please describe any relevant study materials. This could include, for example, stimulus materials used for experiments, questionnaires used for rating studies, training protocols for intervention studies, etc.

For most tasks, a combination of a bullseye and crosshair is used as a fixation target as this combination has been shown to reduce miniature eye movements (Thaler et al., 2013). For task 2 (smooth pursuit), task 3 (microsaccades), and task 5 (pupil dilation grid), a bullseye will be used.

Fixation target:

Bullseye:

Green fixation target which indicates task/trial start:

Additional materials are used in some of the tasks:

- By default, the background is grey, however, for task 5 (luminance), the background color will be changed to different luminances to influence participants' pupil dilation.
- Lastly, a line will be shown for task 7 (head rotation) to indicate to what degree participants should rotate their head.

M14 Study Procedures

Please describe here any relevant information about how the study will be conducted, e.g., the number and timing of measurement time points for longitudinal research, the number of blocks or runs per session of an experiment, laboratory setting, the group size in group testing, the number of training sessions in interventional studies, questionnaire administration for online assessments, etc.

Each participant will be measured individually. Participants will be seated in a room where two PCs will be set up: One will be equipped with the Gazepoint/Tobii eye-tracking devices, the other will be set up with the EyeLink (for the latter, an additional eye-tracker PC is used alongside the stimulus PC). This setup will enable smooth switching between eye-trackers. For both setups, a BenQ XL2430 monitor (refresh rate: 60 Hz,size: 53.13 cm x 29.89 cm, resolution: 1920 x 1080) will be used. For the Gazepoint and Tobii eye-trackers which are mounted to the same monitor above one another (Gazepoint below Tobii), participants will be seated at a distance of 65 cm to the monitor, which is the optimal distance for both eye-tracking devices. For the Eyelink, participants will be seated at a distance of 90 cm to the monitor, and 55 cm to the camera, which is also the optimal distance according to the manual. The room light will be turned off. All eye-tracking devices will record binocularly. The sampling rate of the eye-trackers will be as follows:

- Tobii Pro X3-120: 120 Hz (highest available sampling rate)
- Gazepoint GP3HD: 150 Hz (highest available sampling rate)
- EyeLink 1000+: 1000 Hz

Overall, the experiment will take about 90 minutes. Participants will complete the experiment in two sessions of about 45 minutes each (three out of six blocks will be recorded per session). The break between sessions needs to be at least 15 minutes and up to seven days.

In the first session, age and gender will be inquired. At the start of both sessions, participants will first receive an oral explanation of the upcoming tasks. Informed consent will be obtained, and exclusion criteria will be checked. Then, the experimenter will start the experiment which was programmed in OpenSesame (Mathôt, Schreij & Theeuwes, 2012). During the experiment, the participants' head will be fixed with a chin-forehead rest. For the position tasks, the forehead support will be removed, allowing for head movement into the required position.

An introduction page will be shown, then the first block will start. Between blocks, eye-trackers will be switched, and before each block, the respective eye-tracking device will be calibrated. Tasks will be presented in pseudo-randomized order (see M11). Before each task, a green fixation target is shown to indicate the start of the task. All tasks are described in item M10.

M15 Other information (optional)

Analysis plan

(NOTE: If this varies by hypothesis, repeat analysis plan for each)

AP1 Criteria for post-data collection exclusion of participants, if any

Describe all criteria that will lead to the exclusion of a participant's data (e.g. performance criteria, non-responding in physiological measures, incomplete data). Be as specific as possible.

Pilot data showed that even with an acceptable validation accuracy after the calibration, some data sets showed high measurement errors. To control for this, an outlier correction will be performed (after exclusions on trial level, see AP2). More precisely, three boxplots will be created (one per eye-tracker), in which the overall accuracy values of all individual blocks of all participants will be plotted. For this, not only the accuracy of the large grid (task 1) will be used, but the mean accuracy of all task grids (task 1, 4, and 6). Values identified by the boxplot as extreme outliers (>3xIQR) will be excluded. If one block of a specific participant is excluded, but the other block of the participant for this eye-tracker is not excluded, the present block will be included in the analyses instead of the aggregated data for this participant. If both blocks of one eye-tracker are excluded for a participant, this participant will be excluded altogether and will be replaced.

AP2 Criteria for post-data collection exclusions on trial level (if applicable)

Describe all criteria that will lead to the exclusion of a trial or item (e.g. statistical outliers, response time criteria). Be as specific as possible.

Data will be checked for corruption on trial level. Specifically, after turning the data into standard format, but before any other preprocessing, the percentage of corrupt data within each trial will be inspected. If a trial contains more than 25% corrupt data, this trial will be excluded from further analysis.

AP3 Data preprocessing

Describe all data manipulations that are performed in preparation of the main analyses, e.g. calculation of variables or scales, recoding, any data transformations, preprocessing steps for imaging or physiological data (or refer to publicly accessible standard lab procedure, cf. T12).

Data recorded by each eye-tracking device will be transformed into a standardized format and information about trial and task will be assigned to each data point (see M12).

Samples where both eyes are not valid, which means that both point of gaze measurements are invalid, will be classified as blinks. 100 ms +- each blink will also be classified as blinks.

The remaining corrupt data will be excluded (data outside the monitor frame, data identified as “invalid” by the eye-tracking device, and data for which velocities are >1000 degrees/s or >100.000 degrees/s², as recommended by Nyström & Holmqvist, 2010), also see M8).
Then, saccades, glissades, and fixations will be classified, using the adaptive algorithm suggested by Nyström & Holmqvist (2010).

Data will be aggregated on three levels for the analyses:

- Within blocks: 20% winsorized mean
- Over both sessions: Mean
- Over participants: 20% winsorized mean

AP4 Reliability analysis (if applicable)

Specify the type of scale reliability that will be estimated, whether it is internal consistency (e.g. Cronbach's alpha, omega), test-retest reliability, or some other form (e.g., a confirmatory factor analysis incorporating multiple factors as sources of variance). In a study involving measure development, researchers should specify criteria for removing items from measures a priori (e.g., largest factor loading magnitude, smallest drop in alpha-if-item removed).

AP5 Descriptive statistics

Specify which descriptive statistics will be calculated for which variables. If appropriate, specify which indices of effect size will be used. If descriptive statistics are linked to specific hypotheses, explicitly link the information given here to the respective hypothesis.

The number of events (blinks, saccades, glissades, fixations) as well as their mean duration and amplitude, and the mean accuracy and precision will be inspected 1) separately for each block for each participant, 2) separately for each participant over both sessions recorded with each of the eye-trackers, and 3) over all participants, for each eye-tracker. Furthermore, The number of NA measurements (before preprocessing) will be reported.
The absolute percentage of valid data for point of gaze and pupil measures, as well as the number of identified events (blinks, saccades, glissades, fixations) will be reported for all eye-trackers.

AP6 Statistical models (provide for each hypothesis if varies)

Specify the statistical model (e.g. t test, ANOVA, LMM) that will be used to test each of your hypotheses. Give all necessary information about model specification (e.g., variables, interactions, planned contrasts) and follow-up analyses. Include model selection criteria (e.g., fit indices), corrections for multiple testing, and tests for statistical violations, if applicable. Wherever unclear, describe how effect sizes will be calculated (e.g., for d-values, use the control SD or the pooled SD).

1) Accuracy and precision

1.1) Overall accuracy and accuracy decay (task 1, 4, 6)
Accuracy is defined as the visual angle between the recorded point of gaze and the target location in degree. The accuracy of each block will be calculated as the 20% winsorized mean of each trial's accuracy, only including target fixations (i.e., last fixation of a trial before confirmation), for task 1, 4, and 6 separately. As an overall accuracy estimate, this parameter will be aggregated for each eye-tracker over both sessions (mean) and then on participant level (20% winsorized means). The overall accuracy and the interquartile range (IQR) of each of the three tasks will be reported for each eye-tracker.

Accuracy:
Mixed model A (with data of task 1):
- Dependent variable: Accuracy (aggregated over sessions)
- Fixed effects: Eye-tracker, target, and their interaction
- Random effect: Participant

Accuracy decay:
Mixed model B (with data of task 1, 4, and 6):
- Dependent variable: Accuracy (aggregated over sessions)
- Fixed effects: Eye-tracker, task, and their interaction
- Random effect: Participant

1.2) Overall precision and precision decay (task 1, 4, 6)
As parameters of precision, root mean squared (RMS) between consecutive data points and standard deviation (SD) of point of gaze measurements to the mean fixation location will be calculated for all fixation data of task 1, 4, and 6 separately. Specifically, the standard deviation is calculated as 20% winsorized mean of standard deviations of each fixation of task 1,4, and 6 respectively, per block. RMS and SD are further aggregated over sessions (mean) and over participants (20% winsorized mean). Winsorized mean RMS and SD over all participants will be reported for each eye-tracker.

Precision:
Mixed model A (with data of task 1):
- Dependent variable: RMS (aggregated over sessions)
- Fixed effects: Eye-tracker, target, and their interaction
- Random effect: Participant

Mixed model B (with data of task 1):
- Dependent variable: SD (aggregated over sessions)
- Fixed effects: Eye-tracker, target, and their interaction
- Random effect: Participant

Precision decay:
Mixed model C (with data of task 1, 4, and 6):

- Dependent variable: RMS (aggregated over sessions)
- Fixed effects: Eye-tracker, task, and their interaction
- Random effect: Participant

Mixed model D (with data of task 1, 4, and 6):

- Dependent variable: SD (aggregated over sessions)
- Fixed effects: Eye-tracker, task, and their interaction
- Random effect: Participant

1.3) Accuracy and precision under differing luminance conditions (task 5)
Accuracy and precision will be calculated in a similar fashion to task 1, 4, and 6, for the data of the pupil dilation grid task (task 5), for each luminance level separately. The accuracy and precision estimates per luminance level will be reported.

Mixed model (with data of task 5):

- Dependent variable: Accuracy (aggregated over sessions)
- Fixed effect: Eye-Tracker, target, luminance, eye-tracker * target, eye-tracker * luminance, target * luminance, eye-tracker * target * luminance
- Random effect: Participant

Mixed model (with data of task 5):

- Dependent variable: RMS (aggregated over sessions)
- Fixed effect: Eye-Tracker, target, luminance, eye-tracker * target, eye-tracker * luminance, target * luminance, eye-tracker * target * luminance
- Random effect: Participant

Mixed model (with data of task 5):

- Dependent variable: SD (aggregated over sessions)
- Fixed effect: Eye-Tracker, target, luminance, eye-tracker * target, eye-tracker * luminance, target * luminance, eye-tracker * target * luminance
- Random effect: Participant

For each eye-tracker, t-tests will be computed to compare 1) accuracy, 2) SD, and 3) RMS between task 4 (grid 2) and task 5. Since the pupil dilation grid task (task 5) is shown at a random position in the block (either directly before or after task 4, see M11), it is fairer to compare it with the accuracy of the second grid task (task 4), which is always shown in the middle of the block (always at position 4), instead of comparing it with the accuracy of the first grid (task 1), which is always shown first and probably has the highest accuracy. Specifically, since the pupil dilation grid consists of a subset of five targets out of the large grid, only the corresponding five targets from the large grid will be used for the comparison to increase comparability.

1.4) Accuracy and precision under suboptimal (position) conditions (task 7, 8)
For the tilt position task (task 7) and the turn position task (task 8) separately, the accuracy and precision of the recorded gaze point will be analyzed in a similar way to the overall accuracy (task 1, 4, 6). These accuracy and precision estimates will be reported.

Task 7 (tilt position):
For the tilt position task, because participants will fixate a central fixation target throughout the task, it can be assumed that no new fixations will be identified before confirmation of the trial. Therefore, the 20% winsorized average fixation position 0.5 s before the button press will be used for this calculation.

Mixed model A (with data of task 7):
- Dependent variable: Accuracy (aggregated over sessions)
- Fixed effect: Eye-tracker, line orientation, and their interaction

Random effect: Participant

Mixed model B (with data of task 7):
- Dependent variable: RMS (aggregated over sessions)
- Fixed effect: Eye-tracker, line orientation, and their interaction
- Random effect: Participant

Mixed model C (with data of task 7):
- Dependent variable: SD (aggregated over sessions)
- Fixed effect: Eye-tracker, line orientation, and their interaction
- Random effect: Participant

Task 8 (turn position):
Mixed model D (with data of task 8):
- Dependent variable: Accuracy (aggregated over sessions)
- Fixed effect: Eye-tracker, target, and their interaction
- Random effect: Participant

Mixed model E (with data of task 8):
- Dependent variable: RMS (aggregated over sessions)
- Fixed effect: Eye-tracker, target, and their interaction
- Random effect: Participant

Mixed model F (with data of task 8):
- Dependent variable: SD (aggregated over sessions)
- Fixed effect: Eye-tracker, target, and their interaction
- Random effect: Participant

For each eye-tracker, and for task 7 and 8 separately, t-tests will be computed to compare 1) accuracy, 2) SD, and 3) RMS between task 6 (grid 3) and task 7 / 8 respectively. Since task 7 and 8 are always shown at the end of a block, it is fairer to compare them with the accuracy of the third grid (task 6), which is always shown right before them at the end of the block (always at position 6), instead of comparing them with the accuracy of the large grid (task 1), which is always shown first and probably has the highest accuracy. Specifically, since the position tasks consist of a subset of targets out of the large grid (task 7: central target; task 8: subset of 7 targets), only the corresponding targets from the large grid will be used for the comparison to increase comparability.

2) Point of gaze measurement (X and Y coordinates)

2.1) Mean fixation locations under standard conditions (task 1)
Mixed model A (with data of task 1):
- Dependent variable: Mean fixation location X coordinate (aggregated over sessions)
- Fixed effect: Eye-Tracker, target, and their interaction
- Random effect: Participant

Mixed model B (with data of task 1):

- Dependent variable: Mean fixation location Y coordinate (aggregated over sessions)
- Fixed effect: Eye-Tracker, target, and their interaction
- Random effect: Participant

2.2) Mean fixation locations with varying pupil dilations (task 5)

Mixed model A (with data of task 5):

- Dependent variable: Mean fixation location X coordinate (aggregated over sessions)
- Fixed effect: Eye-Tracker, target, luminance, eye-tracker * target, eye-tracker * luminance, target * luminance, eye-tracker * target * luminance
- Random effect: Participant

Mixed model B (with data of task 5):

- Dependent variable: Mean fixation location Y coordinate (aggregated over sessions)
- Fixed effect: Eye-Tracker, target, luminance, eye-tracker * target, eye-tracker * luminance, target * luminance, eye-tracker * target * luminance
- Random effect: Participant

2.3) Mean fixation locations under suboptimal position conditions (task 7, 8)

Task 7 (tilt position):

For the tilt position task, because participants will fixate a central fixation target throughout the task, it can be assumed that no new fixations will be identified before confirmation of the trial. Therefore, the winsorized average fixation position 0.5 s before the button press will be used for this calculation.

Mixed model A (with data of task 7):

- Dependent variable: Mean fixation location X coordinate (aggregated over sessions)
- Fixed effect: Eye-Tracker, line orientation, and their interaction
- Random effect: Participant

Mixed model B (with data of task 7):

- Dependent variable: Mean fixation location Y coordinate (aggregated over sessions)
- Fixed effect: Eye-Tracker, line orientation, and their interaction
- Random effect: Participant
-

Task 8 (turn position):

Mixed model C (with data of task 8):

- Dependent variable: Mean fixation location X coordinate (aggregated over sessions)
- Fixed effect: Eye-Tracker, target (i.e.. head position), and their interaction
- Random effect: Participant

Mixed model D (with data of task 8):

- Dependent variable: Mean fixation location Y coordinate (aggregated over sessions)
- Fixed effect: Eye-Tracker, target (i.e.. head position), and their interaction
- Random effect: Participant

2.4) Horizontal and vertical bias under suboptimal position conditions (task 7, 8)
Task 7 (tilt position):
The potential influence of the head tilt (task 7) on the data quality will be inspected. Ideally, recorded gaze points should be around (0,0) in this task because participants continue to fixate the central target for the whole task, and it will be assessed how strongly the recorded point of gaze differs from this.

Mixed model A (with data of task 7):
- Dependent variable: Distance between the point of gaze X coordinate and 0 (aggregated over sessions)
- Fixed effect: Eye-tracker, line orientation, interaction
- Random effect: Participant

Mixed model B (with data of task 7):
- Dependent variable: Distance between the point of gaze Y coordinate and 0 (aggregated over sessions)
- Fixed effect: Eye-tracker, line orientation, interaction
- Random effect: Participant

Task 8 (turn position):
Horizontal and vertical bias within the turn task (task 8) will be investigated by computing the difference between recorded coordinate and coordinate of target, separately for x and y coordinates.

Mixed model C (with data of task 8):
- Dependent variable: Distance between the point of gaze X coordinate and the target X coordinate (aggregated over sessions)
- Fixed effect: Eye-tracker, target (i.e.. head position), interaction
- Random effect: Participant

Mixed model D (with data of task 8):
- Dependent variable: Distance between the point of gaze Y coordinate and the target Y coordinate (aggregated over sessions)
- Fixed effect: Eye-tracker, target (i.e.. head position), interaction
- Random effect: Participant

3) Other important eye-tracking parameters

3.1) Smooth pursuit movements (task 2)
Two-phase (hinge) regression models will be calculated for each trial within each participant separately (using the package 'chngpt') to estimate the smooth pursuit onset latency per trial. Specifically, the changepoint until a change in visual angle compared to the center of the screen appears will be calculated as an indicator for smooth pursuit onset latency. This time point will be aggregated over trials (20% winsorized mean), sessions (mean), and participants (20% winsorized mean) to indicate mean smooth pursuit onset latency.
Considering the data between these smooth pursuit onset latencies and the end of each trial, the tracking velocity will be calculated and compared to the corresponding target velocity. *Gain* will be calculated as the ratio between eye velocity and target velocity (which should ideally be 1, see Dowiasch, 2020), and will be aggregated over trials, blocks, and participants (20% winsorized means).

Mixed model A (with data of task 2):
- Dependent variable: Smooth pursuit onset latency (aggregated over sessions)
- Fixed effect: Eye-tracker
- Random effect: Participant

Mixed model B (with data of task 2):
- Dependent variable: Gain (i.e., ratio between eye velocity and target velocity, aggregated over sessions)
- Fixed effect: Eye-tracker
- Random effect: Participant

3.2) Microsaccades (task 3)
The algorithm proposed by Engbert and Mergenthaler (2006) will be run on the subset of data collected in task 3 to identify microsaccades. The mean number of microsaccades (including IQR) will be reported for each eye-tracker.

Mixed model (with data of task 3):
- Dependent variable: Number of identified microsaccades (aggregated over sessions)
- Fixed effect: Eye-tracker
- Random effect: Participant

3.3) Pupil dilation (task 5)
Mean size of both pupils (left and right) will be calculated for each participant and each data point in the pupil dilation grid (task 5). Pupil response will be normalized across participants by dividing the pupil size in each trial by the median of the pupil size recorded while a black screen was presented at the beginning of the task, i.e., the baseline. Thus, the normalized pupil response is the pupil size change relative to the median baseline response in percent. For each luminance level and eye-tracker, the aggregated parameter (20% winsorized mean on the group level) will be reported.

Mixed model (with data of task 5):
- Dependent variable: Normalized pupil response (aggregated over sessions)
- Fixed effect: Luminance level, eye-tracker, and their interaction
- Random effect: Participant

AP7 Inference criteria

Specify the criteria used for inferences (e.g., p values, Bayes factors, effect size measures) and the thresholds for accepting or rejecting your hypotheses. If possible, define a smallest effect size of interest. If inference criteria differ between hypotheses, specify separately for each hypothesis and respective statistical model by explicitly referring to the numbers of the hypotheses. Describe which effect size measures will be reported and how they are calculated.

p-values will be used as inference criteria, with an alpha level of 5%. If multiple analyses are conducted for one parameter (e.g., multiple mixed models are calculated for testing one parameter), alpha will be corrected by using Bonferroni-Holm.

AP8 Exploratory analysis (optional)

Describe any exploratory analyses to be conducted with your data. Include here any planned analyses that are not confirmatory in the sense of being a direct test of one of the specified hypotheses.

AP9 Other information (optional)

All analyses will be conducted in R (R Core Team, 2021), using RStudio (RStudio Team, 2020).

Other information optional

(NOTE: If needed, multiple lines with other information can be included)

O1 Other information (optional)

If there is any additional information that you feel needs to be included in your preregistration, please enter it here. Literature cited, disclosures of any related work such as replications or work that uses the same data, or other context that will be helpful for future readers would be appropriate here.

Beside the technical comparison of data quality between eye-trackers, we will also report the perceived usability of the three eye-tracking devices, to give insight into a broader range of advantages and disadvantages of using one specific device.

Furthermore, the pipeline which was developed to preprocess the data of Gazepoint GP3HD, Tobii Pro X3-120 and EyeLink 1000+ recorded in OpenSesame (PyGaze) into a standardized format will be promoted further, to enhance the standardization within eye-tracking research and to facilitate research using this open software.

References

R1 References

Enter your references below. Use a consistent format (e.g., https://apastyle.apa.org/style-grammar-guidelines/references/examples)

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. Hillsdale, New Jersey: Lawrence Erlbaum Associates. ISBN: 0805802835

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*, 1-23.

Dalmaijer, E. (2014). Is the low-cost EyeTribe eye tracker any good for research? (No. e585v1). *PeerJ PrePrints*.

Dowiasch, S., Wolf, P., & Bremmer, F. (2020). Quantitative comparison of a mobile and a stationary video-based eye-tracker. *Behavior research methods, 52*(2), 667-680.

Ehinger, B. V., Groß, K., Ibs, I., & König, P. (2019). A new comprehensive eye-tracking test
battery concurrently evaluating the Pupil Labs glasses and the EyeLink 1000. *PeerJ* , 7, e7086.

Engbert, R., & Mergenthaler, K. (2006). Microsaccades are triggered by low retinal image slip. Proceedings of the National *Academy of Sciences, 103*(18), 7192-7197.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, *39*(2), 175-191.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods, 41*(4), 1149-1160.

Holmqvist, K. (2017). Common predictors of accuracy, precision and data loss in 12 eye-trackers. In *The 7th Scandinavian Workshop on Eye Tracking*.

Macinnes, J. J., Iqbal, S., Pearson, J., & Johnson, E. N. (2018). Wearable Eye-tracking for Research: Automated dynamic gaze mapping and accuracy/precision comparisons across devices. *bioRxiv* , 299925.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods, 44*(2), 314-324. doi:10.3758/s13428-011-0168-7

Niehorster, D. C., Cornelissen, T. H., Holmqvist, K., Hooge, I. T., & Hessels, R. S. (2018). What to expect from your remote eye-tracker when participants are unrestrained. *Behavior research methods, 50*(1), 213-227.

Nyström, M., Holmqvist, K. An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. *Behavior Research Methods 42*, 188–204 (2010). https://doi.org/10.3758/BRM.42.1.188

R Core Team (2021). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*, Vienna, Austria. URL: https://www.R-project.org/

RStudio Team (2020). *RStudio: Integrated Development for R*. RStudio, PBC, Boston, MA URL http://www.rstudio.com/

Robinson, D. A. (1965). The mechanics of human smooth pursuit eye movement. The *Journal of Physiology, 180*(3), 569-591.

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.

Titz, J., Scholz, A., & Sedlmeier, P. (2018). Comparing eye trackers by correlating their eye-metric data. *Behavior Research Methods, 50*(5), 1853–1863. https://doi.org/10.3758/s13428-017-0954-y

van der Wel, P., & van Steenbergen, H. (2018). Pupil dilation as an index of effort in cognitive control tasks: A review. *Psychonomic bulletin & review, 25*(6), 2005-2015.