UNIVERSITY OF AMSTERDAM

# UvA-DARE (Digital Academic Repository)

## Eye movement benchmark data for smooth-pursuit classification

Korthals, Luke; Visser, Ingmar; Kucharský, Šimon

**DOI**
10.1038/s41597-026-06963-4

**Publication date**
2026

**Document Version**
Final published version

**Published in**
Scientific Data

**License**
CC BY-NC-ND

Link to publication

**Citation for published version (APA):**
Korthals, L., Visser, I., & Kucharský, Š. (2026). Eye movement benchmark data for smooth-pursuit classification. *Scientific Data*, *13*, Article 375. https://doi.org/10.1038/s41597-026-06963-4

**General rights**
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

*UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)*

Download date:05 Sept 2026

scientific data

Check for updates

OPEN

DATA DESCRIPTOR

# Eye movement benchmark data for smooth-pursuit classification

Luke Korthals[1]✉, Ingmar Visser[1,3] & Šimon Kucharský[2,3]

**Analysis of eye tracking data often requires accurate classification of eye movement events. Human experts and classification algorithms often confuse episodes of fixations (fixating stationary targets) and smooth pursuits (fixating moving targets) because their feature characteristics overlap. To foster the development of better classification algorithms, we created a benchmark data set that does not rely on human annotation as the gold standard. It consists of almost four hours of eye movements. Ten participants fixated different targets designed to induce saccades, fixations, and smooth pursuits. Plausible benchmark labels were established by designing stimuli that prevent fixations and smooth pursuits to co-occur, and separating them from saccades by their velocity. Here we make available both the raw data and offer a convenient way for preprocessing and assigning plausible benchmark labels in the form of a companion package in Python. We encourage researchers to utilize them for feature engineering, and to train, validate, and benchmark their algorithms.**

## Background & Summary

Eye movement research relies on accurately classifying episodes of fixations, smooth pursuits, saccades, and other eye movements (see Lappi, 2016 for an overview[1]). Fixations actively maintain the gaze on or near stationary targets, while smooth pursuits are active attempts at "fixating" moving targets by matching their velocity[1]. In terms of the samples (consecutive gaze coordinates) collected during eye tracking, both fixations and smooth pursuits are characterized by low sample-to-sample velocities[1–3]. In contrast, saccades are ballistic eye movements (high sample-to-sample velocity) that rapidly shift the gaze between locations[1]. While fixations and saccades are readily distinguished based on their velocity, many state-of-the-art event classification algorithms fail to reliably distinguish smooth pursuits from fixations and sometimes confuse smooth pursuits and saccades[3–10]. Similarly, human annotators find smooth pursuits difficult to classify and frequently disagree on their classification[2]. Furthermore, it is questionable whether expert annotations constitute a gold standard[2,11]. This lack of accurate automatic smooth pursuit classification is a serious shortcoming because, among other applications, smooth pursuits are used to investigate schizophrenia[12], traumatic brain injury[13], and neurodegenerative diseases[14,15].

Real-world eye movement classification involves complex factors like head, body, and binocular dynamics, and naturalistic stimuli[1]. In line with this, some of the most prominent datasets for training, evaluating and comparing classification algorithms are based on eye movements in response to naturalistic stimuli and human annotation (see https://michaeldorr.de/smoothpursuit/ for an overview). However, here we address the core difficulty of reliably distinguishing smooth pursuits from other eye movements by deliberately collecting data under highly controlled conditions (fixed-head, monocular eye movements, moving dots). This procedure enabled us to avoid manual annotation and instead establish plausible benchmark labels for fixations, saccades, and smooth pursuits only based on experimental manipulation and dynamic velocity thresholds. Our objective was to collect data that makes it easy to identify statistical features that can improve the classification of smooth pursuits. This approach may yield a reliable classification foundation that can then be generalized and validated in more complex, naturalistic datasets such as GazeCom[16].

This dataset consists of nearly four hours of eye movement recordings collected at 1000 Hz from ten university students using a fixed-head paradigm. Participants were instructed to focus on small moving circles without moving their head. These circles, moved according to different patterns, speeds, and trajectories and were designed to evoke fixations, saccades, and smooth pursuits. As the head and body remained stationary, gaze and target positions were collected in a screen-relative (world) coordinate system. Importantly, each trial was limited

[1]University of Amsterdam, Department of Psychology, Amsterdam, 1018 WS, The Netherlands. [2]Technical University Dortmund, Department of Statistics, Dortmund, 44227, Germany. [3]These authors jointly supervised this work: Ingmar Visser, Šimon Kucharský ✉e-mail: l.korthals@uva.nl

to include maximally one type of slow eye movement events (i.e., fixations OR smooth pursuits) and one type of fast eye movements (i.e., saccades). This way we aimed to prevent fixations and smooth pursuits – which are frequently confused during annotation – from co-occurring; thus, enabling us to label eye movements based on a combination of the experimental design and adaptive velocity thresholds, without relying on human annotation. Additionally, this separation of fixations and smooth pursuits should make it easier to identify novel statistical features (feature engineering) that are better at distinguishing between these eye movements than velocity or other commonly used features. In addition to sharing the raw eye movement recordings, we also provide the full experimental code and target trajectories, enabling researchers to collect comparable data and investigate if incorporating target information into algorithms can improve classification performance. We think that this dataset complements existing datasets well and we encourage other researchers to use it for feature engineering, to train, validate, and benchmark their algorithms, and to investigate data-driven approaches other than dynamic velocity thresholds for assigning benchmark labels.

Data were randomly assigned to training (n = 7) and test sets (n = 3) and are available on the Open Science Framework (OSF; https://osf.io/zx7hc/)[17]. The OSF repository contains raw tracking data, participant demographics, target locations, calibration and validation performance, and a video showcasing the stimuli. We also offer a companion package on GitHub (https://github.com/lukekorthals/pursuing-smooth-pursuits-data). This repository contains the code and instructions for downloading, preprocessing, annotating, and plotting the data. A second GitHub repository (https://github.com/lukekorthals/pursuing-smooth-pursuits-experiment) contains the experimental code and a script for interactively visualizing the different stimuli.

## Methods

The study was approved by the Ethics Review Board of the Faculty of Social and Behavioural Sciences at the University of Amsterdam, reference number FMG-2710. Participants were recruited via direct email and provided written, informed consent prior to taking part in the study. They agreed to the participation and sharing of their data for research purposes. The retained and published data does not personally identify the participants. The consent procedure included an explanation of the study aims, procedures, voluntary nature of participation, right to withdraw, and data-handling practices.

**Participants.** Twelve psychology students from the University of Amsterdam participated as a favor to the researchers. Two participants could not be calibrated because the eye tracker tracked their pupil jerkily and frequently lost it completely; thus, we collected no data for them. The remaining ten participants were between 22 and 27 years old (M = 24.2). Most identified as female (7 female, 2 male, 1 other), and had brown eyes (brown = 7, blue = 3). Eight participants were nearsighted; half of them used vision correction during data collection (glasses = 1, contact lenses = 3, no correction = 4). The remaining two participants had no known vision impairments. All participants indicated that they could clearly discern the stimuli.

**Technical Setup.** Participants stabilized their head using a chin rest, positioning their eyes 70 cm from the center of a widescreen monitor (61.4 × 36.5 cm; 2560 × 1440 pixels) operating at a 144 Hz refresh rate. Monocular eye movements of the right pupil were collected at 1000 Hz with an EyeLink 1000 Plus[18] using ellipse tracking without drift correction. The EyeLink was mounted on the table below the center of the monitor.

**Calibration and Validation.** All calibration and validation results are accessible in the data repository. Participants self-calibrated using the manufacturer-provided nine-point calibration by focusing on circles and pressing the `Enter` key on a keyboard. Each calibration attempt was validated using the same nine-point procedure. We tried to achieve fair or good validation according to the SR Research manual[19] for all participants and had them repeat the procedure multiple times if it indicated poor calibration. Eight participants achieved fair or good calibration quality, with a mean average error of 0.45 degrees (°) and a mean maximum error of 1.18°. Two participants only achieved poor calibration (average/max errors = 1.08°/3.16° and 0.50°/2.13° respectively). Degrees here denote the angular distance on the retina between the true target and the measured gaze position – sometimes referred to as degrees of visual angle (dva)[20]. Throughout this manuscript, we use degrees to express not only calibration error but also stimulus size (the angle subtended at the eye) and stimulus velocity in degrees per seconds (°/s).

Calibration and validation were supposed to be repeated between trials whenever participants removed their head from the chin rest during a break or if the eye tracker lost their pupil. This happened once for two participants each. They achieved good and fair validation during recalibration (average/max errors = 0.8°/1.22° and 0.79°/1.65° respectively). After investigating the logs, we realized that participant `21db28aa` took a two-minute break after trial 81 but was not recalibrated, causing the gaze in all subsequent trials to be severely misaligned with the targets (see "Target Alignment" in Table 1). Consequently, we suggest removing trials 82 to 144 when analyzing data for this participant based on the `trial_number` column.

**Stimuli.** Stimuli were created based on self-testing by two of the researchers (LK and SK). During each trial a small grey circle (#808080; radius = 0.2°) was presented on a black background (#000000). Participants were instructed to focus on these circles (targets) without moving their head and ideally without blinking. Circles moved across the screen and their movement pattern, speed, and trajectory were varied between trials.

*Movement Pattern.* Circles followed one of three patterns to evoke specific eye movements. Each pattern was designed to elicit only one type of fast (i.e., saccades) and one type of slow eye movements (i.e., fixations or smooth-pursuits) per trial. "Moving circles" moved consistently at target speed to elicit smooth pursuits (slow eye movements). "Jumping circles" jumped along multiple fixation locations to induce series of fixations (slow eye movements) and saccades (fast eye movements). Fixation locations were spaced evenly according to target

| Participant | Trials | AVG Validation Error (dva) | MAX Validation Error (dva) | Missing Data (%) | Fixation Precision (dva) | Target Alignment (cosine similarity) |
|---|---|---|---|---|---|---|
| 06b8d2d3 | 1-144 | 1.08 | 3.16 | 1.33 | 0.11 | 0.91 |
| 21db28aa | 1-81 | 0.50 | 2.13 | 1.16 | 0.14 | 0.95 |
| | 82-144 | — | — | 1.43 | 0.05 | −0.21 |
| 68471e16 | 1-15 | 0.53 | 1.30 | 0.52 | 0.08 | 0.96 |
| | 16-144 | 0.80 | 1.22 | 0.48 | 0.09 | 0.91 |
| 6cde27b5 | 1-144 | 0.49 | 1.28 | 2.2 | 0.09 | 0.90 |
| 7d248f8f | 1-144 | 0.30 | 0.84 | 0.12 | 0.13 | 0.94 |
| 88878fe6 | 1-144 | 0.45 | 1.64 | 2.29 | 0.09 | 0.91 |
| cf910821 | 1-144 | 0.31 | 0.94 | 0.05 | 0.1 | 0.94 |
| d2c0afa4 | 1-144 | 0.57 | 0.91 | 0.24 | 0.07 | 0.93 |
| eb8c52e6 | 1-15 | 0.69 | 1.56 | 0.68 | 0.23 | 0.87 |
| | 16-144 | 0.79 | 1.65 | 2.2 | 0.12 | 0.84 |
| f7bb2338 | 1-144 | 0.25 | 0.97 | 3.46 | 0.17 | 0.93 |

**Table 1.** Measures indicate good data quality for all trials apart from 82–144 of participant `21db28aa`.

speed and circles stood still for one second at each location. "Back-and-forth circles" moved consistently at target speed while jumping back and forth along the axis of their trajectory at one second intervals to produce smooth pursuits (slow eye movements) interspersed by (catch-up) saccades (fast eye movements).

*Movement Distance and Target Speed.* We selected the movement distance and speed based on self-testing to satisfy the following requirements.

- **Participants should accurately track targets at all speeds without moving their head**. This was to ensure that the collected eye movements could be reliably classified relative to the world coordinate system.
- **Participants should comfortably complete each trial without or with minimal blinking**. This was necessary to ensure the collected gaze trace clearly represents the intended eye movements and maximizes the data completeness for classification algorithms.
- **We wanted to collect data for very slow and very fast smooth pursuits**. We aimed for this because statistical features for distinguishing smooth pursuits from fixations and saccades may vary across different pursuit speeds.
- **Smooth pursuits at all speeds should be clear without intermittent fixations or catch-up saccades**. Avoiding fixations is required to classify eye movements based on experimental manipulations and dynamic velocity thresholds. We also wanted to avoid catch-up saccades during "moving circle" trials to ensure that pure smooth pursuit features may be extracted from our data. In contrast, "back-and-forth circle" trials were designed to include catch-up saccades.
- **Trials had to last at least one second long**. This ensured that "jumping circle" and "back-and-forth circle" trials would include at least a single saccade.

We first selected a movement distance of 20° because it falls within normal eye movement ranges[21] and felt comfortable without moving the head during self-testing. Then we selected target speeds to cover slow, medium, and fast smooth pursuit speeds. Maximum smooth pursuit speeds without catch-up saccades are reported to be around 30°/s[1]. At this speed, targets would clear the distance in less than one second, which would be too fast, as it would not allow for even a single saccade. To our knowledge, no minimum speeds for smooth pursuits are reported in the literature. Therefore, we piloted different speeds for "moving circle" trials. We found it difficult to smoothly track targets slower than 1°/s and faster than 6°/s. This is consistent with Drewes *et al.* (2018)[22], who found that smooth pursuits slower than 1°/s included fixations, and faster than 6°/s included saccades. Based on self-testing we selected speeds of 1°/s, 3°/s, and 6°/s to represent slow, medium and fast smooth pursuits respectively. At these speeds, targets took 20 s, 6.66 s, and 3.33 s respectively to clear 20°. With these settings we could comfortably track targets without moving our heads and with minimal blinking. They also ensured that smooth pursuits would be collected at varying speeds and "jumping circle" and "back-and-forth circle" trials would include at least three saccades. Therefore, these settings satisfied all our requirements.

Inadvertently, the code for converting degrees to pixels did not correctly account for the monitor dimensions. As a result, targets did not move the intended 20° at speeds of 1°/s, 3°/s, and 6°/s. Instead, horizontal movements covered approximately 19.29° at 0.96°/s, 2.89°/s, and 5.78°/s, vertical movements covered approximately 20.38° at 1.02°/s, 3.06°/s, and 6.11°/s, and diagonal movements covered approximately 28.06° at 1.4°/s, 4.21°/s, and 8.41°/s. The actual distance and speed for each trial are included in the data (`actual_distance` and `actual_speed` columns). For transparency, we share the functions that were used during the experiment (PREPROCESSOR._PIX2DEG_FAULTY(), PREPROCESSOR._DEG2PIX_FAULTY() and the functions we used to calculate the actual distances and speeds (PREPROCESSOR._PIX2DEG(), PREPROCESSOR._DEG2PIX()) here: https://github.com/lukekorthals/pursuing-smooth-pursuits-data/blob/main/eyemovement_data/preprocessor.py.

*Target Trajectory.* Targets moved along eight straight trajectories; specifically, cardinal and ordinal directions (north ↑, south ↓, west ←, east →, north-west ↖, north-east ↗, south-west ↙, south-east ↘). North and south represent up- and downwards movements, while east, and west represent right-, or leftwards movements. We selected this variety of horizontal, vertical, and diagonal movements because they are each controlled by different muscle pairs[23]. Additionally, we avoided circular movements because plots reported by Drewes *et al.* (2018)[22] indicate they are more difficult to track.

*Task Battery.* We presented every participant with each pattern-speed-trajectory combination twice for a total of 144 trials (i.e., pattern [3] × speed [3] × trajectory [8] × block [2]). Trials took 20 s, 6.66 s, and 3.33 s for targets moving approximately 1°/s, 3°/s, and 6°/s respectively. Therefore, about 24 minutes of eye movements were collected for each participant and almost four hours were collected in total. Note that slow targets were overrepresented in this dataset because their trials took longer. This makes this dataset particularly useful for studying how to distinguish slow smooth pursuits from fixations. Trials were organized in two blocks. Within each block, trials were ordered from slow to fast while movement pattern and trajectory were randomized. We did not randomize speed because adjusting to slower targets after tracking multiple fast targets appeared difficult during pilot testing.

**Data Acquisition Procedure.** Data for all participants was collected by one of the researchers (LK) in June 2023 in an eye tracking lab at the University of Amsterdam. Upon entering the lab, the experimenter informed participants about the procedure, and they signed participation forms. Participants sat in front of the screen, rested their head in a chin rest, and calibrated the eye tracker using nine-point calibration. Calibration was repeated multiple times in an attempt to achieve fair or good validation. After calibration the experimenter instructed participants to focus on all targets without moving their head and to avoid blinking as much as possible. Participants completed a tutorial trial (moving circle, 2°/s, east →) and could ask questions. Afterwards, participants started each trial by pressing the `Space` bar. Between trials, they could take as many breaks as they wanted. Before the first trial, the experimenter emphasized the importance of collecting clear data and encouraged them to take frequent breaks. To keep participants focused, the experimenter told them they were doing well and encouraged them to keep it up after the first trial and intermittently throughout data collection. Additionally, to ensure the quality of collected eye movements the experimenter constantly monitored if the eye tracker was working and if participants were accurately following the targets. If a participant blinked frequently or lost track of the target, the experimenter reminded them to focus and take breaks. If participants removed their head from the chin rest or the eye tracker lost their pupil, calibration was repeated before the next trial. After the last trial, the experimenter thanked and debriefed participants. Each session took between 45 and 60 minutes, depending on the time spent for breaks and calibration.

## Data Records

The OSF repository (https://osf.io/zx7hc/) includes an mp4 video showcasing some stimuli and an R script that is used to extract relevant data from raw data files. It also contains two folders named "raw" and "clean", each with subfolders "train" and "test". These contain data folders for seven and three participants respectively. For "raw" data, each participant folder holds five raw data files and for "clean" data, each participant folder holds six csv files[17].

**Files under "raw":**

- **<id>.EDF**: The raw EyeLink data.
- **<id>.asc**: ASCII conversion of the EDF file.
- **<id>_participant.csv**: Participant age, sex, and eye condition.
- **<id>_targets.csv**: Target coordinates for each trial, experiment and trial time to match with gaze data.
- **<id>_trials.csv**: Movement pattern, distance, and speed for each trial.

**Files under "clean":**

- **<id>_blinks.csv**: EyeLink blink data extracted from the ASCII file inside "raw".
- **<id>_eyetracker.csv**: EyeLink settings extracted from the ASCII file inside "raw".
- **<id>_gaze.csv**: EyeLink gaze coordinates extracted from the ASCII file inside "raw".
- **<id>_participant.csv**: A copy of the raw `<id>_participant.csv` without the `experiment_name` column.
- **<id>_targets.csv**: A copy of the raw `<id>_targets.csv` without the `experiment_name` column.
- **<id>_trials.csv**: A copy of the raw `<id>_trials.csv` without the `experiment_name` column and convenient trajectory labels.

## Technical Validation

Data quality calculations can be replicated by following the instructions inside the `data_quality.ipynb` script. We looked at five indicators to assess data quality for each participant:

- **Average validation error:** Reported by the EyeLink. Indicates how accurately the pupil position was tracked. Should ideally be below 0.5 degrees for adults with normal vision[19].

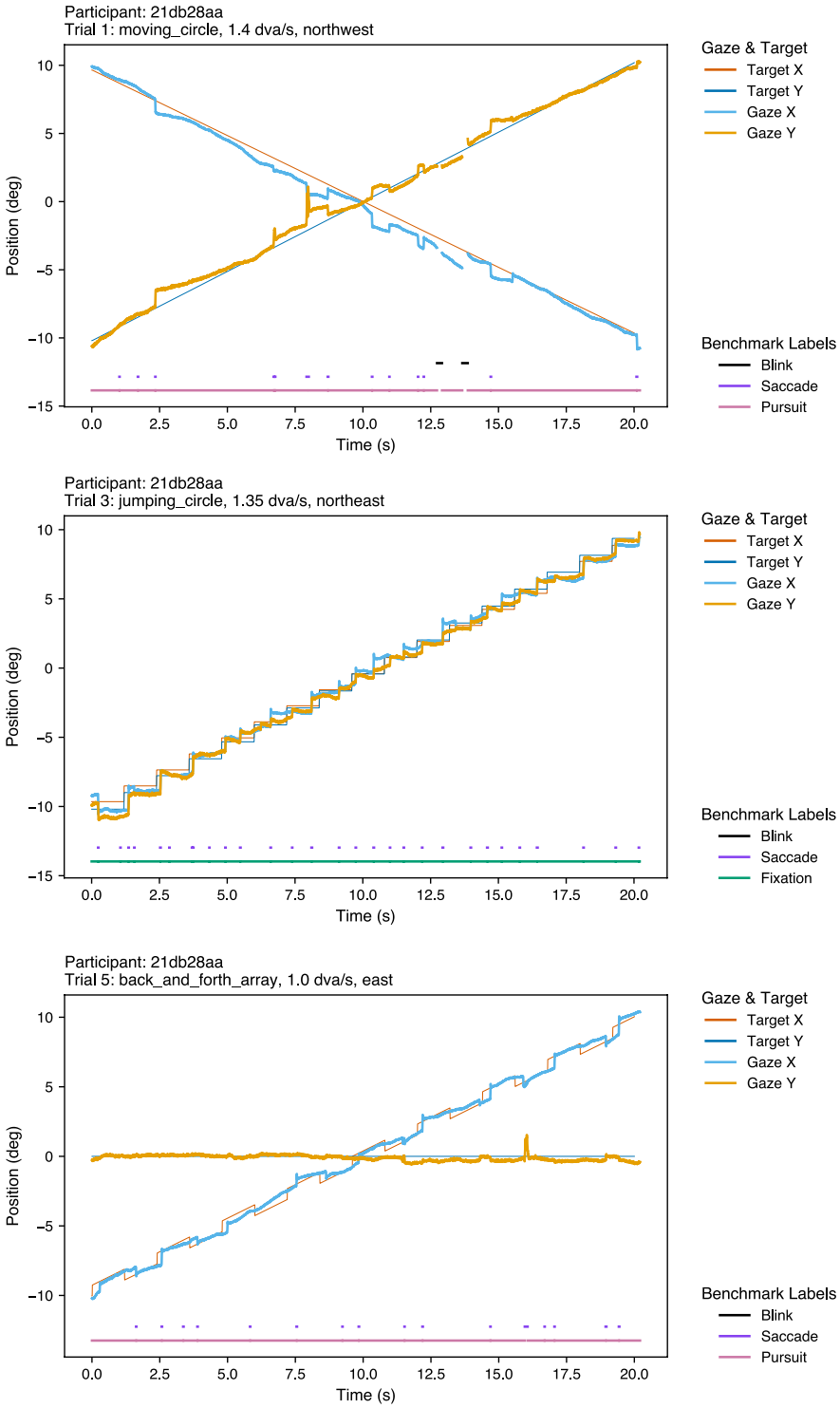


**Fig. 1** These plots show one trial for each of the three movement patterns. They were created with the PARTICIPANT.PLOT_TRIAL() function and include participant and trial information, gaze and target positions, and the annotated benchmark labels. We used https://davidmathlogic.com/colorblind/ for colorblind friendly colors.

- **Maximum validation error:** Reported by the EyeLink. Indicates how accurately the pupil position was tracked. Should ideally be below 1.0 degree for adults with normal vision[19].
- **Missing data:** Percent of gaze samples with *NA* values. Indicates how prevalent blinks and tracking problems are in the data. Should be as low as possible.

- **Fixation precision:** Root Mean Squared Error (RMSE) of gaze samples relative to the centroid of each classified fixation group in "jumping circle" trials. Reflects the spatial dispersion of gaze within fixations. Lower values indicate more stable, tightly clustered fixations on stationary targets.
- **Target alignment:** Cosine similarity between gaze and target trajectories. Indicates how diligently participants followed targets with their eyes. Should be as close to 1 as possible.

Table 1 shows the results for all participants. We calculated the five indicators for all trials that followed (re) calibration of a given participant. Overall, the indicators suggest good data quality with the exception of trials 82 to 144 for participant `21db28aa`, who took a two-minute break after trial 81 and was not recalibrated.

## Usage Notes

In addition to the raw data, we share a companion Python package. Follow the installation instructions on GitHub (https://github.com/lukekorthals/pursuing-smooth-pursuits-data) and run the `quickstart.ipynb` script. This will make it easy to download, preprocess, and annotate the data; and includes basic plotting functions (see Fig. 1). In the following, we provide a brief overview of the preprocessing and classification protocols.

**Cleaning.** Raw data is cleaned by calling the CLEAN_RAW_DATA() function. Cleaning refers to extracting all relevant data from raw EyeLink and PsychoPy files, storing it as csv, removing unnecessary columns (e.g., `experiment_name`), setting convenient column names (e.g., `xp` becomes `gaze_x`), and setting convenient trajectory values (e.g., "ver_up" becomes "north").

**Preprocessing.** Clean data is preprocessed by calling the PARTICIPANT.PREPROCESS_CLEAN_DATA() function. By default, the PREPROCESSOR.ORIGINALPREPROCESSOR(PREPROCESSOR) class is used to implement the procedure outlined below. You can adjust some parameters like the window size for blink removal or create your own subclass of the PREPROCESSOR.PREPROCESSOR() to implement your own preprocessing protocol.

*Correcting Target positions for back-and-forth circle trials.* Due to a bug, target locations collected during "back-and-forth circle" trials were shifted along their jumping positions (e.g., target at position A but position B collected). Due to another bug, the starting position of these targets was reversed for participants `06b8d2d3`, `6cde27b5` and `21db28aa` (e.g., target started at position B instead of position A). Consequently, the correct target positions are recalculated according to trial time, target speed, and trajectory.

*Transforming gaze and target positions.* Eye and target positions were initially collected in pixels; here they are converted to degrees and rounded to three decimals. Time measurements of eye and target positions are converted to seconds and rounded to three decimals. Eye positions were collected relative to the center of the screen, while target positions were collected relative to the top left corner. Therefore, eye positions are rescaled to match target positions. While eye positions were collected at 1000 Hz, target positions could only be collected at the monitor refresh rate of 144 Hz; consequently, missing target positions are filled with previous target positions.

*Removing EyeLink delays.* To ensure that the EyeLink was tracking during each entire trial, tracking started 100 ms before trial onset and ended 100 ms after trial offset. Consequently, all gaze samples before target onset and after target offset are removed.

*Removing blinks.* Blinks are characterized by scattered, fast samples when the eye tracker loses the pupil[24]. These artifacts could affect the velocity threshold approach we propose to distinguish saccades from fixations or smooth pursuits because of their high speed. Therefore, all samples the EyeLink classified as blinks were set to *NA*. Additionally, in line with the approach by Lüken *et al.*[3], all samples 50 ms before and after the blinks classified by the EyeLink are also set to *NA* as they consist of very fast samples that may interfere with dynamic velocity thresholding. Like other parameters, users can adjust the window size for blink removal when using our companion package.

**Classifying Plausible Benchmark Labels.** As indicated above, stimuli were designed such that each trial should only include saccades (fast eye movement), and fixations OR smooth pursuits (slow eye movements). While we cannot rule out that (stationary) fixations occurred during smooth pursuit trials, because participants may have occasionally "zoned out", we are confident that they generally followed the instructions because we provided continuous encouragement and monitored the eye-tracker signal. Evidence in the literature indicates that fast and slow eye movements can be distinguished based on their velocity[2–5,8,10,16,25–27]. This should be especially the case for the current dataset because the fastest target speed for smooth pursuits is around 6°/s while saccades typically show velocities upwards of 100°/s[1]. As long as participants diligently followed the targets with their eyes, "moving circle" trials should include smooth pursuits and (possibly) unintended saccades. "Back-and-forth circle" trials should include smooth pursuits and saccades. "Jumping circle" trials should include fixations and saccades. Based on these assumptions we developed the following protocol to create benchmark labels through dynamic velocity thresholds.

Note that this dataset can also be used without establishing benchmark labels in purely data driven endeavors. Additionally, we encourage other researchers to improve our protocol for establishing benchmark labels and/or developing different protocols altogether. For convenience, the following protocol is included in our companion Python package and can be used by calling the PARTICIPANT.CLASSIFY_PREPROCESSED_DATA() function as described in `quickstart.ipynb`.

*Calculating dynamic thresholds.* A dynamic velocity threshold to distinguish slow from fast eye movements is calculated for each trial. For a given trial the velocity is first calculated as the Euclidean distance divided by time. Then the threshold is determined by taking the median velocity and adding 1.5 times the 75th percentile velocity (see Equation (1)). This approach assumes that most gaze movements were slow — either due to stationary targets or consistently moving targets designed to elicit fixations and smooth pursuits. In contrast, saccades (either induced by abrupt target jumps or arising from participant inattention or blinks) were expected to produce brief bursts of higher gaze velocity, exceeding the computed threshold. We landed on the median, 75th percentile, and a scaling factor of 1.5 based on visual inspection of plots from arbitrarily selected trials of participants `f7bb2338` and `6cde27b5`.

$$\text{Threshold} = \text{Median Velocity} + 1.5 \times \text{75th Percentile Velocity} \quad (1)$$

*Distinguishing fast and slow eye movements.* All samples with velocities exceeding the threshold are classified as saccades. The remaining samples are classified as fixations for "jumping circle" trials, or as smooth pursuits for "moving circle" and "back-and-forth circle" trials.

*Relabeling short events.* Lappi (2016)[1] reports that the shortest microsaccades have durations of 10 ms. We also use 10 ms as a very conservative estimate for minimum fixation and smooth pursuit durations to allow for short events. Fixations and smooth pursuits shorter than 10 ms are reclassified as saccades. Finally, any remaining saccades shorter than 10 ms are relabeled as fixations for jumping circle trials or smooth pursuits for moving circle and back-and-forth circle trials; thus, concluding classification.

## Data availability

All data is available on OSF (https://osf.io/zx7hc/). The companion package (https://github.com/lukekorthals/pursuing-smooth-pursuits-data) makes it easy to download and use it.

## Code availability

We programmed the experiment in PYTHON 3.8.16[28]. Specifically, we used PSYCHOPY 2023.1.3[29] for the experimental process, the Eyelink Developers Kit with PYLINK 0.3.3[30] to communicate with the eye tracker, and NUMPY 2.2.3[31] and PANDAS 2.2.3[32] to calculate how stimuli moved across the screen. The code used to collect the data is available on GitHub (https://github.com/lukekorthals/pursuing-smooth-pursuits-experiment). The companion package for downloading, preprocessing and annotating data is also available on GitHub (https://github.com/lukekorthals/pursuing-smooth-pursuits-data).

Received: 29 August 2025; Accepted: 25 February 2026;
Published online: 16 March 2026

## References

1. Lappi, O. Eye movements in the wild: Oculomotor control, gaze behavior & frames of reference. *Neuroscience and Biobehavioral Reviews* **69**, 49–68, https://doi.org/10.1016/j.neubiorev.2016.06.006 (2016).
2. Andersson, R., Larsson, L., Holmqvist, K., Stridh, M. & Nyström, M. One algorithm to rule them all? An evaluation and discussion of ten eye movement event-detection algorithms. *Behavior Research Methods* **49**, 616–637, https://doi.org/10.3758/s13428-016-0738-9 (2017).
3. Lüken, M., Kucharský, Š. & Visser, I. Characterising eye movement events with an unsupervised hidden Markov model. *Journal of Eye Movement Research* **15**, https://doi.org/10.16910/jemr.15.1.4 (2022).
4. Komogortsev, O. V. & Karpov, A. Automated classification and scoring of smooth pursuit eye movements in the presence of fixations and saccades. *Behavior Research Methods* **45**, 203–215, https://doi.org/10.3758/s13428-012-0234-9 (2013).
5. Larsson, L., Nyström, M., Andersson, R. & Stridh, M. Detection of fixations and smooth pursuit movements in high-speed eye-tracking data. *Biomedical Signal Processing and Control* **18**, 145–152, https://doi.org/10.1016/j.bspc.2014.12.008 (2015).
6. Santini, T., Fuhl, W., Kübler, T. & Kasneci, E. Bayesian identification of fixations, saccades, and smooth pursuits. In *Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications*, ETRA '16, 163–170 https://doi.org/10.1145/2857491.2857512 (Association for Computing Machinery, New York, NY, USA, 2016) .
7. Pekkanen, J. & Lappi, O. A new and general approach to signal denoising and eye movement classification based on segmented linear regression. *Scientific Reports* **7**, 17726, https://doi.org/10.1038/s41598-017-17983-x (2017).
8. Startsev, M., Agtzidis, I. & Dorr, M. 1D CNN with BLSTM for automated classification of fixations, saccades, and smooth pursuits. *Behavior Research Methods* **51**, 556–572, https://doi.org/10.3758/s13428-018-1144-2 (2019).
9. Startsev, M., Agtzidis, I. & Dorr, M. Characterizing and automatically detecting smooth pursuit in a large-scale ground-truth data set of dynamic natural scenes. *Journal of Vision* **19**, 10, https://doi.org/10.1167/19.14.10 (2019).
10. Dar, A. H., Wagner, A. S. & Hanke, M. REMoDNaV: robust eye-movement classification for dynamic stimulation. *Behavior Research Methods* **53**, 399–414, https://doi.org/10.3758/s13428-020-01428-x (2021).
11. Hooge, I. T. C., Niehorster, D. C., Nyström, M., Andersson, R. & Hessels, R. S. Is human classification by experienced untrained observers a gold standard in fixation detection? *Behavior Research Methods* **50**, 1864–1881, https://doi.org/10.3758/s13428-017-0955-x (2018).
12. O'Driscoll, G. A. & Callahan, B. L. Smooth pursuit in schizophrenia: A meta-analytic review of research since 1993. *Brain and Cognition* **68**, 359–370, https://doi.org/10.1016/j.bandc.2008.08.023 (2008).
13. Hunfalvay, M. *et al.* Vertical smooth pursuit as a diagnostic marker of traumatic brain injury. *Concussion* **5**, CNC69, https://doi.org/10.2217/cnc-2019-0013 (2020).
14. Tao, L. *et al.* Eye tracking metrics to screen and assess cognitive impairment in patients with neurological disorders. *Neurological Sciences* **41**, 1697–1704, https://doi.org/10.1007/s10072-020-04310-y (2020).
15. Martinez-Conde, S., Macknik, S. L. & Hubel, D. H. The role of fixational eye movements in visual perception. *Nature Reviews Neuroscience* **5**, 229–240, https://doi.org/10.1038/nrn1348 (2004).
16. Dorr, M., Martinetz, T., Gegenfurtner, K. R. & Barth, E. Variability of eye movements when viewing dynamic natural scenes. *Journal of Vision* **10**, 28, https://doi.org/10.1167/10.10.28 (2010).
17. Korthals, L., Visser, I. & Kucharský, Š. Eye-movement benchmark data for smooth-pursuit classification, https://doi.org/10.17605/OSF.IO/ZX7HC (2024).

18. SR Research. EyeLink 1000 Plus. https://www.sr-research.com/eyelink-1000-plus/. [Apparatus and software].
19. SR Research. What determines whether a validation is "GOOD", "FAIR" or "POOR"? https://www.sr-research.com/support/showthread.php?tid=244 (2020).
20. Holmqvist, K. *et al*. *Eye tracking: A comprehensive guide to methods and measures* (Oxford University Press, United States, 2011).
21. Lee, W. J., Kim, J. H., Shin, Y. U., Hwang, S. & Lim, H. W. Differences in eye movement range based on age and gaze direction. *Eye* **33**, 1145–1151, https://doi.org/10.1038/s41433-019-0376-4 (2019).
22. Drewes, H., Khamis, M. & Alt, F. Smooth Pursuit Target Speeds and Trajectories. In *Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia*, MUM '18, 139–146, https://doi.org/10.1145/3282894.3282913 (Association for Computing Machinery, New York, NY, USA, 2018).
23. Pinkhardt, E. H. & Kassubek, J. Ocular motor abnormalities in Parkinsonian syndromes. *Parkinsonism & Related Disorders* **17**, 223–230, https://doi.org/10.1016/j.parkreldis.2010.08.004 (2011).
24. Hershman, R., Henik, A. & Cohen, N. A novel blink detection method based on pupillometry noise. *Behavior Research Methods* **50**, 107–114, https://doi.org/10.3758/s13428-017-1008-1 (2018).
25. Komogortsev, O. V., Gobert, D. V., Jayarathna, S., Koh, D. H. & Gowda, S. M. Standardization of automated analyses of oculomotor fixation and saccadic behaviors. *IEEE Transactions on Biomedical Engineering* **57**, 2635–2645, https://doi.org/10.1109/TBME.2010.2057429 (2010).
26. Larsson, L., Nyström, M. & Stridh, M. Detection of saccades and postsaccadic oscillations in the presence of smooth pursuit. *IEEE Transactions on Biomedical Engineering* **60**, 2484–2493, https://doi.org/10.1109/TBME.2013.2258918 (2013).
27. Salvucci, D. D. & Goldberg, J. H. Identifying fixations and saccades in eye-tracking protocols. In *Proceedings of the 2000 symposium on Eye tracking research & applications*, ETRA '00, 71–78, https://doi.org/10.1145/355017.355028 (Association for Computing Machinery, New York, NY, USA, 2000).
28. Van Rossum, G. & Drake Jr, F. L. *Python reference manual* (Centrum voor Wiskunde en Informatica Amsterdam, 1995).
29. Peirce, J. *et al*. PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods* **51**, 195–203, https://doi.org/10.3758/s13428-018-01193-y (2019).
30. SR Research. Getting Started with Python and PyLink. https://www.sr-research.com/support/thread-48.html (2020).
31. Harris, C. R. *et al*. Array programming with NumPy. *Nature* **585**, 357–362, https://doi.org/10.1038/s41586-020-2649-2 (2020).
32. McKinney, W. Data structures for statistical computing in Python. In Walt, S. v. d. & Millman, J. (eds.) *Proceedings of the 9th Python in Science Conference*, 56 – 61, https://doi.org/10.25080/Majora-92bf1922-00a (2010).


## Author contributions

I.V.: Conceptualization, Methodology, Project administration, Supervision, and Writing - review & editing. L.K.: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing - original draft, and Writing - review & editing. Š.K.: Conceptualization, Formal analysis, Methodology, Project administration, Software, Supervision, Validation, and Writing - review & editing.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to L.K.

**Reprints and permissions information** is available at www.nature.com/reprints.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.


**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2026