eyeris: A flexible, extensible, and reproducible pupillometry preprocessing framework in R

Shawn T. Schwartz[1,2,3,4,5*], Haopei Yang[1,2], Alice M. Xue[1,2,3,5], Mingjian He[1,2,6]

[1]Stanford Department of Psychology, Stanford University, United States
[2]Stanford Memory Laboratory, Stanford, United States
[3]Stanford University Wu Tsai Neurosciences Institute, Stanford, United States
[4]Stanford University Wu Tsai Human Performance Alliance, Stanford, United States
[5]Stanford University Center for Mind, Brain, Computation and Technology, Stanford, United States
[6]Department of Anesthesiology, Perioperative, and Pain Medicine, Stanford Medicine, United States
*For correspondence: stschwartz@stanford.edu (S.T.S.)

Author ORCIDs

Shawn T Schwartz (0000-0001-6444-8451)
Haopei Yang (0000-0002-1207-3245)
Alice M Xue (0009-0006-3176-8786)
Mingjian He (0000-0002-6688-8693)

Competing Interests

We declare that there are no conflicts of interest with respect to the authorship or publication of this article.

Abstract

Pupillometry provides a non-invasive window into the mind and brain, particularly as a psychophysiological readout of autonomic and cognitive processes like arousal, attention, stress, and emotional states. Pupillometry research lacks a robust, standardized framework for data preprocessing, whereas in functional magnetic resonance imaging and electroencephalography, researchers have converged on tools such as *fMRIPrep*, *EEGLAB* and *MNE-Python*; these tools are considered the gold standard in the field. Many established pupillometry preprocessing packages and workflows fall short of serving the goal of enhancing reproducibility, especially since most existing solutions lack designs based on Findability, Accessibility, Interoperability, and Reusability (FAIR) principles. To promote FAIR and open science practices for pupillometry research, we developed *eyeris*, a complete pupillometry preprocessing suite designed to be intuitive, modular, performant, and extensible (*https://github.com/shawntz/eyeris*). Out-of-the-box, *eyeris* provides a recommended preprocessing workflow and considers signal processing best practices for tonic and phasic pupillometry. Moreover, *eyeris* further enables open and reproducible science workflows, as well as quality control workflows by following a well-established file management schema and generating interactive output reports for both record keeping/sharing and quality assurance of preprocessed pupil data prior to formal analysis. Taken together, *eyeris* provides a robust all-in-one transparent and adaptive solution for high-fidelity pupillometry preprocessing with the aim of further improving reproducibility in pupillometry research.

Impact Statement

Pupillometry research currently lacks a standardized, integrated preprocessing framework comparable to tools widely adopted in EEG and fMRI research. We introduce *eyeris*, an open-source R package that fills this gap through a modular, transparent pipeline with signal processing best practices, interactive diagnostic reports for quality control, and scalable database storage. *eyeris* advances pupillometry methods by promoting reproducible, FAIR-compliant workflows accessible to researchers at all levels of programming expertise.

Keywords

pupillometry, eye tracking, preprocessing, reproducibility, signal processing, open-source software

1 Introduction

The measurement of pupil size and its relation to cognition dates back to the late 1800s with investigation of the pupillary light reflex (*Haab, 1886*). More than a century later, pupillometry underwent a present-day (re)discovery, establishing pupil size as an integrated psychophysiological readout of cognitive processes such as attention. More specifically, pupillometry offers a non-invasive window into the human mind and brain, and is sensitive to environmental (i.e., luminance, focal distance) and cognitive (i.e., alerting, orienting, attention, executive functioning) changes (*Huang and Clewett, 2024; Strauch, 2024; Strauch et al., 2022*).

The breadth of cognitive processes that have been associated with pupil size dynamics is expansive, including – but not limited to – emotional arousal (*Bradley et al., 2008; Oliva and Anikin, 2018*), mental effort (*Kahneman and Beatty, 1966; Kramer et al., 2013*), fatigue/task performance (*Aston-Jones and Cohen, 2005a; Eldar et al., 2013; Knapen et al., 2016; McGinley et al., 2015; Murphy et al., 2011*), response preparation (*Einhäuser et al., 2008; Reimer et al., 2014*), uncertainty/decision-making (*Gee et al., 2014; Jepma and Nieuwenhuis, 2011; Kawaguchi et al., 2018*), prediction error (*Alamia et al., 2019*), working memory (*Keene et al., 2022; Robison and Unsworth, 2019; Unsworth and Robison, 2015, 2017a,b*), and episodic memory (*Clewett et al., 2020; Madore et al., 2020; Madore and Wagner, 2022; Miller et al., 2019; Miller and Unsworth, 2020; Robison et al., 2022; Schwartz et al., 2025*).[1]

As highlighted above, the majority of pupillometry studies use pupil size as an assay of attentional engagement or examine the impact of cognitive processes on pupil size (*Kristjansson et al., 2009; Madore et al., 2020; van den Brink et al., 2016*). Pupil size is not only linked with various cognitive processes, but also has specific neural substrates, such that pupil diameter dynamics are oftentimes interpreted as a non-invasive (indirect) surrogate of locus coeruleus-noradrenergic (LC–NA) neuromodulatory activity (*He et al., 2020; Joshi et al., 2016; cf. Megemont et al., 2022*). In humans, pupil diameter has been shown to covary with functional magnetic resonance imaging (fMRI) blood-oxygenation level–dependent (i.e., BOLD) activity in LC (*Murphy et al., 2014*). More recently, fMRI has been used to identify neural networks in the human brain associated with pupil diameter dynamics in healthy young and older adults and revealed decreases in functional connectivity within a putative pupil diameter network with healthy aging (*Wu et al., 2025*). In most studies, analyses of these pupil data are conducted offline following data collection, although there are also real-time applications providing causal evidence confirming the use of moment-to-moment pupil dynamics as an assay of attentional engagement (*Keene et al., 2022; Meissner et al., 2024; Schwartz et al., 2024, 2025*). Additionally, it is worth noting that given LC's central role in regulating autonomic function and arousal (*Poe et al., 2020; Samuels and Szabadi, 2008*) and that LC is the first brain region to develop tau pathology in Alzheimer's disease (*Betts et al., 2019; Braak and Del Tredici, 2011; Holper et al., 2022;*

[1] See *Fink et al. (2024)* for a detailed table further outlining documented associations between cognitive processes and pupil size.

Spires-Jones and Hyman, 2014), pupillometry is also being used as a noninvasive window into cognitive aging (*El Haj, 2024; Kim et al., 2024; Piquado et al., 2010*).

Consequently, collecting high-quality pupil data and appropriately analyzing these data are critical for advancing both theories of cognition as well as translational applications of pupillometry. Moreover, data preprocessing is a fundamental prerequisite to high quality pupillometry data analysis (*Fink et al., 2024; Hershman et al., 2024; Kret and Sjak-Shie, 2019; Laeng and Mathôt, 2024; Mathôt and Vilotijević, 2023; Steinhauer et al., 2022; van Rij et al., 2019*).[2] In pupillometry research, experimenters deal with fewer time series, at most two for binocular recordings, compared to electroencephalography (EEG) or fMRI, but many of the same principles apply: handling missing data, excluding artifacts, and smoothing out physiologically irrelevant noise, to name a few.

Numerous researchers have contributed excellent software packages to the community over the years to aide others in preparing and analyzing pupillometry data without the need to implement core preprocessing functions themselves (e.g., *Forbes, 2020; Geller et al., 2020; Hershman et al., 2019; Kinley and Levy, 2021; Tsukahara, 2020*). However, most of the existing tools only provide rigid preprocessing specifications, often have been factored out of particular analysis projects, and/or do not empower users to be deliberate in preprocessing choices and to compare visualizations of preprocessed pupil data to that of raw pupil data.

Thus, a critical gap in current pupillometry research is the absence of an integrated, standardized, and adaptable preprocessing pipeline – a point emphasized by *Mathôt and Vilotijević (2023)*, who note:

> There is no commonly agreed-upon workflow for conducting cognitive-pupillometry experiments: design criteria, preprocessing steps, and statistical analyses all differ vastly between studies. Some attempts at standardization have been made, but these are not easily applicable to cognitive pupillometry [...] Furthermore, guidelines are often conceptual rather than specific implementations, and the lack of concrete examples and code makes it difficult for researchers to incorporate these guidelines into their own workflow. (p. 3055)

In contrast, tools developed for other neuroimaging modalities over the past couple of decades, notably *fMRIPrep* for fMRI data (*Esteban et al., 2020, 2019*), have defined new standards for transparent and reproducible data preprocessing workflows. While widely used software packages such as *EEGLAB* (*Delorme and Makeig, 2004*) and *MNE-Python* (*Gramfort et al., 2013*) for M/EEG provide extensive preprocessing functionality, they operate primarily as flexible toolboxes, leaving decisions about preprocessing steps, their order, and parameter settings to individual researchers. More recently, efforts to build standardized preprocessing pipelines on top of these toolboxes – such as *PREP* (*Bigdely-Shamlo et al., 2015*) and *Automagic* (*Pedroni et al., 2019*) for EEG, or *MNE-BIDS-Pipeline*

[2] Throughout this manuscript, "preprocessing" means the specific prescription of steps necessary to derive transformed pupil data ready for visualization and analysis from the raw pupil data recorded from a participant using an eye tracking device.

(*https://mne.tools/mne-bids-pipeline/stable/*) for M/EEG – reflect a growing recognition of the need for more prescribed, reproducible preprocessing workflows.

However, developing and sustaining robust preprocessing pipelines is not trivial. Many scientists with strong domain expertise lack formal training in computer science or software engineering (*Chen et al., 2025; Segal, 2008*), a likely contributing factor to the inadvertent adoption of common programming "anti-patterns" when writing analytic code (*Wilson et al., 2014*). In fact, nonreproducibility in science often results from some combination of poor programming practices (*Hatton, 1997; Hatton and Roberts, 1994; Merali, 2010; Miller, 2006*), suboptimal and inattentive code and data management and sharing practices (*Frank et al., 2025; Gorgolewski and Poldrack, 2016; Kovacs et al., 2024*), and/or insufficient transparency about the specific details underlying each step taken in between to transform raw data into a high quality format appropriate for inferential statistics (*Poldrack et al., 2019*). Over and above these pitfalls is the "garden of forking paths" problem and the culprit of Type I (i.e., false positive) errors in extant literature (*Gelman and Loken, 2013; Poldrack et al., 2017; Simmons et al., 2011*).

In the context of psychological science, flexibility in data analysis and reporting has been implicated as a prominent contributor to the replication and reproducibility crisis (*Hardwicke et al., 2021, 2018, 2022; Ioannidis, 2005; Open Science Collaboration, 2015*). Likewise, the garden of forking paths for psychophysiological data is a slippery slope: when researchers are afforded more preprocessing pathways and analytic degrees of freedom, even small adjustments to the way data are preprocessed (e.g., digital filter design, handling of missing data, artifact removal) can impact the outcomes and interpretations a researcher may draw from their data (*Šoški'c et al., 2024*).

To fill this gap in robust preprocessing pipelines for pupillometry data, we developed *eyeris*, an R package dedicated to providing unified standards for pupillometry preprocessing (*Schwartz, 2025a*). With *eyeris*, we not only intend to provide users with an extensible toolkit for conducting robust pupillometry preprocessing, but critically aim to contribute to standardizing best practices in pupillometry research (especially for cognitive pupillometry; see *Mathôt (2018a)* for a review discussing pupil responses, the neural pathways subserving these pupillary responses, and their relationship to high-level cognition). We should note that we are not the first group to put forth a recommended set of preprocessing guidelines for cognitive pupillometry. Here, we have built upon the comprehensive guide to cognitive pupillometry preprocessing put forth by *Mathôt and Vilotijević* (*2023*); their stated aim is to provide "guidelines [that] are meant to be illustrative, rather than prescriptive; that is, we show how things could be done – rather than how they should be done – in order to implement a workflow that is both easy to implement and that meets contemporary standards for good scientific practice" (p. 3056). In contrast, we adopt a prescriptive, rather than merely *illustrative*, approach to pupillometry preprocessing workflows.

Here, we argue that a standardized cognitive pupillometry preprocessing workflow is essential to limit the number of forking paths available to researchers during analysis. In other words, without consistent preprocessing, studies leveraging pupillometry to advance key theories in psychological science could be hindered by high degrees of freedom and vulnerable to unintended mistakes during preprocessing. As such, our accompanying R package is designed with "glass box" principles (see *Esteban et al., 2019; Poldrack et al., 2019*), thereby equipping both novice and experienced pupillometry researchers with key tools needed to appropriately transform raw pupillometry data into a format well-suited for downstream analysis, visualization, and statistical inference. Embracing a carefully designed pipeline is highly beneficial, as it enables researchers to focus on hypothesis testing with minimal worry about whether the data were preprocessed correctly. In this era of open-source software, it is also important to allow experienced users to inspect and modify steps in a preprocessing pipeline with minimal hurdles.

eyeris strongly encourages users to follow a prescribed set of steps in a particular order and with carefully selected default parameters. The overarching aim is to avoid unnecessary preprocessing decisions that may lead to issues when made with incomplete knowledge of some of the nuances of signal processing. At the same time, *eyeris* provides transparent access and configurability to all preprocessing steps in a user-friendly fashion, which is supported by its modular, glass-box design (***Figure 1***, ***Box 1***). This design approach greatly reduces the burden of usage by removing the need to develop lengthy scripts for individual use cases (*Varoquaux, 2016*).

Box 1. Example usage of the *glassbox* pipeline wrapper

```
library(eyeris)

path_to_demo_dataset <- eyelink_asc_demo_dataset()

cleandata <- glassbox(file = path_to_demo_dataset)
#> [  OK  ] - Running eyeris::load_asc()
#> [  OK  ] - Running eyeris::deblink()
#> [  OK  ] - Running eyeris::detransient()
#> [  OK  ] - Running eyeris::interpolate()
#> [  OK  ] - Running eyeris::lpfilt()
#> [ SKIP ] - Skipping eyeris::detrend()
#> [  OK  ] - Running eyeris::zscore()

plot(cleandata)
#> ! Plotting block 1 from possible blocks: 1
```

2 Method

2.1 eyeris R package

The *eyeris* R package is actively maintained and available on both CRAN and GitHub (see *Code Availability*). ***Box 2*** describes the three recommended methods for installing *eyeris*.

Box 2. *eyeris* R package installation options

```
# (a) to install the latest stable release from CRAN:
install.packages("eyeris")

# (b) using Posit's pak approach to install packages interactively:
# install.packages("pak")
pak::pak("eyeris")

# (c) to install the bleeding edge development version:
# install.packages("devtools")
devtools::install_github("shawntz/eyeris")
```

2.2 Entry-point function: glassbox()

The primary entry point for preprocessing pupillometry data in *eyeris* is the glassbox() function, which offers an opinionated, fully transparent preprocessing pipeline. This "glass-box" design contrasts with "black-box" workflows by allowing (and encouraging) users to understand and evaluate the consequences of each processing step on their data. Within *eyeris*, the glassbox() function completely streamlines the preprocessing workflow, enabling researchers to seamlessly transform raw pupil data into a high-quality, standardized format ready for analysis and visualization.

We designed glassbox() to operate in both *interactive* and *headless* modes. In the interactive mode, users are well-positioned to evaluate pre- and post-processing plots of their pupil time series data following the execution of each pipeline step, with prompts to continue or modify parameters as needed. In the headless mode, users are permitted to automate preprocessing of pupil datasets of any size, including large-scale pupillometry datasets on cloud computing clusters[3], using recommended default parameters without manual intervention.

Users are permitted to override the default parameters of any preprocessing step within glassbox() by passing a named list of arguments to the specific function (e.g., deblink = list(extend = 100)) as arguments to the glassbox() function. Taken together, *eyeris*' flexibility out-of-the-box facilitates real-time parameter optimization and empowers researchers to tailor the pipeline to their specific data characteristics and research questions in a transparent fashion.

[3] See *https://github.com/shawntz/clamp* for a complete implementation of *eyeris* in headless mode for deployment on high-throughput cloud compute clusters (using the Slurm Workload Manager).

2.3 Sequential preprocessing steps of the glass-box pipeline implementation and usage

The glassbox() pipeline executes a fully transparent prescribed sequence of pupillometry preprocessing steps, which are also available as individual modular functions (see ***Box 1***). Although users are able to build their own pipelines using the individual modules, we strongly recommend that users invoke the glassbox() pipeline function, given each step is critical and must be performed in a specific order to avoid unexpected data modifications that could compromise downstream analyses. Below, we demonstrate the sequence of core methods that comprise the glassbox() function to showcase both its flexibility and its syntactic simplicity.

2.3.1 Loading and parsing data: eyeris::load_asc()

This initial step builds upon the *eyelinker::read.asc()* function and parses SR Research EyeLink *.asc* files[4], extracting raw time series data, event messages, blink events, and various metadata from the file header. The load_asc() function within *eyeris* also supports automatic handling of multiple recording segments (i.e., blocks) within a single .asc file and offers options for gracefully handling binocular data out-of-the-box (e.g., averaging left and right eye pupil sizes, or processing them independently). Note, it is critical that users load their data using eyeris::load_asc() given there are a number of configuration steps *eyeris* performs to ensure data can be seamlessly passed through each pipeline module.

2.3.2 Blink artifact removal: eyeris::deblink()

The deblink() function removes blink-related artifacts by NA-padding missing-data periods and their surrounding samples in the pupil time series. Each period of missing data is extended both backward and forward in time by a user-configurable duration (defaulting to 50 ms) to mask the partial-occlusion artifacts that typically flank eye blinks. The *extend* parameter accepts either a single value for symmetric padding or a two-element vector for asymmetric padding (e.g., *c(backward, forward)*). The appropriate *extend* duration may vary across experimental setups or participant populations and can be optimized upon inspection of the diagnostic reports generated by *eyeris*.

2.3.3 Transient artifact removal: eyeris::detransient()

The detransient() function identifies and removes non-physiological pupil samples – such as transient spikes caused by unstable tracking or head movements – by applying a speed-based threshold. The

[4] At this time, *eyeris* only supports EyeLink-acquired pupillometry datasets, given its prevalence in this research space. However, we acknowledge that accessibility to EyeLink hardware is limited due to equipment costs, and thus are now working on building in a porting mechanism for other eye tracking data files. To achieve this goal, we will first create a generic solution that can take a CSV file of recording data and will specify a mapping between key column names in that file with the expected labels utilized by *eyeris*. A more ambitious goal is native support for proprietary formats from other trackers; this step will require external contributions and requests from users with access to these data formats that we could use to implement the compatibility layer, which we will evaluate on a case-by-case basis via submissions to the issues tab on GitHub (see *Code availability* for more information on how to access the source code and/or contribute directly to the development of *eyeris*).

function computes an approximate pupil change velocity using finite differences and flags samples whose speed exceeds a threshold derived from the Median Absolute Deviation (MAD; *Arachchige and Prendergast, 2024; Hampel, 1968, 1974; Leys, 2013; Rousseeuw and Croux, 1993; Rousseeuw et al., 1986*) of the velocity distribution. Specifically, the threshold is defined as the median speed plus *n* times the MAD value, where the constant *n* defaults to 16 (*Tsukahara, 2020*). For edge cases where the computed MAD or median speed is very small, users may supply a manual threshold via the *mad_thresh* parameter to adjust sensitivity.

2.3.4 Linear interpolation: eyeris::interpolate()

The interpolate() function fills in missing pupil samples resulting from the preceding blink removal and artifact detection steps, recovering a contiguous time trace. It applies linear interpolation via the *na.approx()* function from the *zoo* package (*Zeileis & Grothendieck, 2005*), with edge NAs replaced by the nearest available values. Linear interpolation is well-suited to pupillometry given the relatively slow dynamics of pupil size fluctuations; however, users are strongly encouraged to set cutoff thresholds to exclude trials or periods in which a substantial proportion (e.g., 50%) of raw data are missing.

2.3.5 Low-pass filtering: eyeris::lpfilt()

The lpfilt() function smooths the pupil time series by applying a low-pass Butterworth filter, with the default passband cutoff at 4 Hz and stopband at 8 Hz. The filter order is automatically determined by the *buttord()* function from the *gsignal* package (*van Boxtel, 2021*) to satisfy the specified passband/stopband frequencies and ripple/attenuation constraints. The resulting filter uses a second-order sections representation for numerical stability. Users may modify the filter design parameters (*wp*, *ws*, *rp*, *rs*) and inspect the frequency response by setting *plot_freqz = TRUE*.

2.3.6 Downsampling: eyeris::downsample()

The downsample() function reduces the sampling frequency of the pupil time series to a user-specified target rate (*target_fs*) while preserving the original temporal dynamics of the signal. Unlike binning, downsampling retains individual data points after first applying an automatically designed anti-aliasing low-pass filter to prevent spectral aliasing. As a safeguard, the function raises an error if the resulting passband frequency would fall below 4 Hz, protecting frequency content relevant to pupillary responses. Users can optionally visualize the anti-aliasing filter's frequency response by setting *plot_freqz = TRUE*.

2.3.7 Binning: eyeris::bin()

The bin() function provides an alternative approach to reducing temporal resolution by dividing the time series into equal-duration intervals and computing a summary statistic – either the mean (default) or median – within each bin. The number of bins per second is controlled by the

bins_per_second parameter, and the resulting timepoints are centered within each interval. Although binning is commonly used in pupillometry research, users should be aware that averaging within bins can attenuate rapid pupillary dynamics; the choice between downsampling and binning should therefore be guided by the specific research question.

2.3.8 Detrending: eyeris::detrend()

The detrend() function removes slow linear drifts from the pupil time series by fitting a linear model of pupil size as a function of time and returning the residuals. This step can be useful for removing gradual drifts such as the commonly reported "time-on-task effect", but is disabled by default in *eyeris* for two reasons: first, trial-level baseline correction often renders detrending unnecessary (see *Mathôt et al., 2018b* for a detailed discussion of the tradeoffs between subtractive and divisive baseline correction approaches), and second, linear detrending is an imperfect correction – slow trends in pupil size are often nonlinear, and applying a linear detrend can introduce new biases to the signal (*Strauch et al., 2015*). Moreover, slow trends may reflect physiologically meaningful tonic arousal levels that should be preserved depending on the research question (see ***Section 3.1.1*** for a more detailed discussion). Users may enable this step within *glassbox()* by setting *detrend = TRUE* when the removal of slow linear trends is appropriate for their research question.

2.3.9 Z-scoring: eyeris::zscore()

The zscore() function normalizes the pupil time series to have a mean of 0 and a standard deviation of 1 by mean-centering and scaling, analogous to *base::scale()*. This normalization facilitates trial-level and between-subjects analyses, as the arbitrary units of pupil size do not scale consistently across participants or recording sessions. Because *eyeris* is designed to preprocess one block of data per participant at a time, each block is z-scored independently; when preprocessed data are later merged, all blocks will have been separately scaled. Z-scoring is enabled by default and can be disabled by setting *zscore = FALSE* within *glassbox()*.

2.4 Additional functions for epoching, exporting, and visualizing pupil data

In addition to the core preprocessing steps described above, *eyeris* provides built-in methods to conveniently extract tidy trial-based epochs with optional baseline correction, export BIDS-style derivatives and diagnostic HTML reports, and interactively browse diagnostic plots for tens of thousands of trials at each preprocessing step.

2.4.1 Epoching and baseline correction: eyeris::epoch()

The epoch() function is intended to be used as the final preprocessing step. It creates data epochs of either fixed or dynamic durations with respect to user-supplied event message strings and time limits, and includes an intuitive metadata parsing feature where additional trial data embedded within event

messages can be automatically identified and joined into the resulting epoched data frames. The function supports three main event-modes: (*i*) a single string representing the event message to center epoch extraction around, using specified limits (or, if no limits are provided, extracting data between each subsequent occurrence of the same event); (*ii*) a vector containing both start and end event message strings, where limits are ignored and the duration of each trial epoch is determined by the number of samples between each matched start–end pair; and (*iii*) a list of two data frames that manually specify start/end event timestamp–message pairs for extraction from the raw time series, with explicit block number specification.

For event-modes (*i*) and (*ii*), the event message string must conform to a standardized protocol so that *eyeris* knows how to find events and optionally parse included metadata. Users have two primary choices: (a) specify a string followed by a * wildcard expression (e.g., *PROBE_START**), which matches any messages beginning with that prefix; or (b) specify a string using the *eyeris* curly-brace syntax (e.g., *PROBE_{type}_{trial}*), which matches messages following that structure and automatically generates additional metadata columns (e.g., *type* and *trial*) in the tidy output. The function also supports optional baseline correction via the *calc_baseline* and *apply_baseline* parameters, with both subtractive and divisive correction types and flexible specification of the baseline period through either fixed time windows or event-delimited intervals (see ***Box 3*** for additional details).

Box 3. Tips & Tricks

Extracting pupil data from trial epochs with varying durations.

Users may also be interested in extracting the pupil time series for each stimulus/probe from start-to-finish. To illustrate, if each probe was displayed for a variable duration — perhaps because the participant's response designates the end of that trial — then the user cannot rely on a fixed-duration epoch extraction method. To demonstrate the power and flexibility of our implementation for this case in *eyeris*, take, for instance, an expansion of the researcher's event message tags for the probe onsets we discussed in the main text above and suppose that the researcher also included event message tags to signal the end of any given epoch, such as: *PROBE_END_32_PIANO*. With *eyeris*, the user can specify their starting boundary to be *PROBE_START_{trial}_{stimulus}* and ending boundary to be *PROBE_END_{trial}_{stimulus}*, which tells *eyeris* to grab all samples between each of these boundaries, and additionally, to create new metadata columns in the tidy output data frame containing the third and fourth extracted values in between the underscore separator with column names of *trial* and *stimulus*, respectively. See *Function: epoch()* for additional details.

2.4.2 Exporting BIDS-like derivatives: eyeris::bidsify()

The bidsify() function provides a structured way to save preprocessed pupil data in a BIDS-like directory structure. It exports epoched data as well as the raw pupil time series, and formats directory and file name structures based on user-supplied metadata (participant ID, session number, task name, and run number). For multi-block recordings contained within a single *.asc* file, blocks are automatically numbered as separate runs. The function also supports merging epochs into a single output file, merging runs across blocks, and optionally generating interactive HTML diagnostic reports.

These reports include epoch-by-epoch diagnostic plots that can be grouped by a user-specified column variable (defaulting to the matched event label). In the future, we intend for this function to save data in an official BIDS format for eye tracking data (see the *proposal currently under review*); however, at this time, the function takes a BIDS-inspired approach to organizing output files (see ***Box 4***).

2.4.3 Diagnostic visualization: eyeris::plot()

The plot() function is an S3 plotting method for objects of class *eyeris*. It plots a single-panel time series for a subset of the pupil data at each preprocessing step, providing a simple method for qualitatively assessing the consequences of the preprocessing recipe and its parameters on the raw pupillary signal. Users can control which preprocessing steps to visualize (defaulting to all), the number and duration of random preview epochs, and whether to display accompanying histograms of pupil samples at each step via the *plot_distributions* parameter. A fixed random seed can be set to ensure the same preview epochs are displayed across repeated calls, which is especially useful when iteratively adjusting parameters within and across workflow steps. For multi-block recordings, the *block* parameter specifies which block to plot.

2.5 Database storage and analysis

In addition to the BIDS-like file exports described further in ***Section 3.1.4***, *eyeris* provides an integrated database storage backend for centralized management of preprocessed pupillometry data. Database storage is activated by setting *db_enabled = TRUE* within the *bidsify()* function, which creates or connects to a DuckDB (*Raasveldt & Mühleisen, 2019*) database file (*.eyerisdb*) in the derivatives directory. All data types generated by the preprocessing pipeline – including time series, events, blinks, epochs, epoch summaries, and confounds – are automatically written to the database alongside any CSV exports. For large-scale deployments where storage efficiency and I/O performance are priorities, CSV output can be disabled entirely by setting *csv_enabled = FALSE*.

The *eyeris* database API provides a suite of functions for connecting to, querying, and extracting data from the database without requiring SQL expertise. The *eyeris_db_summary()* function provides a quick overview of the database contents, including available subjects, sessions, tasks, and row counts per data type. The *eyeris_db_collect()* function serves as the primary extraction interface, enabling researchers to aggregate data across multiple subjects in a single call with flexible filtering by subject, session, task, epoch label, and eye. For advanced use cases, direct SQL queries can be executed through the underlying DBI connection obtained via *eyeris_db_connect()*.

To support data sharing and distribution of large databases, *eyeris* includes chunked export functionality. The *eyeris_db_to_chunked_files()* function processes data incrementally (defaulting to 1 million rows per chunk) and writes the output to CSV or Parquet files, automatically splitting files that

exceed a configurable size threshold (defaulting to 50 MB). For formal data distribution, *eyeris_db_split_for_sharing()* partitions the database into portable chunks that can be reconstructed by collaborators using *eyeris_db_reconstruct_from_chunks()*. The Parquet format is also supported, offering 50–80% file size reduction compared to CSV, faster read performance, and preserved numeric precision.

Box 4. BIDS-like derivatives data structure

```
eyeris/
|- derivatives/
  |- sub-001/
        |- eye/
        |  |- sub-001_task-ret_run-01_desc-timeseries_pupil.csv
        |  |- sub-001_task-ret_run-01_epoch-probe_desc-preproc_pupil.csv
        |- source/
        |  |- figures/
        |        |- run-01/
        |        |- epoch_probe/
        |        |  |- run-01_PROBE_22_1.png
        |        |  |- run-01_PROBE_22_2.png
        |        |  |- run-01_PROBE_22_3.png
        |        |  |- ...
        |        |- run-01_fig-1_desc-histogram.jpg
        |        |- run-01_fig-1_desc-timeseries.jpg
        |- sub-001_epoch-probe_run-01.html
        |- sub-001.html
```

3 Results

3.1 Key Features & Innovations

Here, we introduce *eyeris* with the goal of making pupillometry preprocessing more efficient, reproducible, and conducive for reliable inferences as well as providing researchers greater confidence when analyzing their pupillometry data.

3.1.1 Common pupillometry preprocessing steps performed in a prescribed order.

Pupillometry time series require a relatively small number of preprocessing steps compared to multivariate data (such as EEG and fMRI); yet, there is a more defined set of issues to be addressed when preprocessing pupil size data. Critically, these preprocessing steps must be completed in a certain order to avoid unexpected modifications of the data in a way that could compromise downstream analyses. In this section, we describe some of the most common preprocessing steps for pupillometry data, which are implemented in the current version of *eyeris*. While additional steps might be added in the future, some of the basic principles covered here will be applicable in general.

First, preprocessing of pupil sizes needs to handle artifactual or missing data due to eye blinking, movement, or simply incorrect pupil detection and measurement by the eye tracker. Most existing commercial eye trackers, including SR Research EyeLink[5], have built-in blink detection methods that will annotate blink periods in exported data. One can easily recognize blink periods when the pupil size becomes 0/NaN for a brief period of time. However, it is important to recognize that there are often blink-related artifactual changes in pupil size surrounding an eye blink, which can be due to the eyelid rapidly covering the pupil area detected by the camera, resulting in rapidly diminishing pupil size before becoming zero. In contrast, as the eyelid opens, there is a brief period of rapidly increasing pupil size when the camera detects a partial pupil. Occasionally, eye trackers may incorrectly lock onto other structures in the camera's field of view, such as eyebrows and glasses frames, as they may become more probable pupil targets compared to partially occluded pupils during blinking. This can also result in fluctuations in recorded pupil sizes. As such, *eyeris* removes blink periods using the deblink() function with an *extend* duration (defaults to 50 ms) bleeding both backward and forward in time to mask neighboring time periods that may contain such artifacts. Note, the *extend* duration parameter may need to be adjusted for different experimental setups, certain groups of participants, or even for individual participants, which can be determined on a case-by-case basis upon inspection of visual diagnostic reports.

While commercial eye trackers can mark blink periods for easy exclusion, there are also artifactual pupil size measures that could occur in recordings due to unstable pupil tracking or eye/head movements. These non-blink disruptions of pupillometry measurement often result in transient and abrupt changes in recorded pupil sizes, commonly visible as spikes in the time series. *eyeris* detects and removes such transient artifacts using the detransient() function. Under the hood, detransient() computes an approximate pupil change velocity by taking the first-order difference divided by sampling interval and detects outliers using the MAD procedure (*Arachchige and Prendergast, 2024; Hampel, 1968, 1974; Leys, 2013; Rousseeuw and Croux, 1993; Rousseeuw et al., 1986*). In our experience, most artifacts in pupillometry data can be captured using the simple detransient() method, thereby excluding non-physiological measures. However, noisier recordings with different eye tracker hardware may require more elaborate artifact detection algorithms, such as to remove sustained periods of outlying pupil sizes instead of velocity. Users are strongly encouraged to expand on the function given their data and contribute to *eyeris* with additional methods to remove artifact periods (see ***Box 7*** and ***Appendix*** for more information about extensibility and developing custom extensions in *eyeris*).

Once data from blink and artifact periods are removed, the next step in pupillometry preprocessing is to interpolate over these periods to recover a contiguous time trace of pupil sizes (*Mathôt et al., 2013*). While not strictly required for certain types of pupillometry analysis – such as averaging across trial periods for a univariate measure – many recent analytic approaches for pupillometry data leverage

[5] https://www.sr-research.com

each sampled time point to uncover temporal dynamics in changes of pupil sizes (*Clewett et al., 2020; Johansson et al., 2018; Steinhauer et al., 2004; Urai et al., 2017; Van Slooten et al., 2018; Verney et al., 2004; Widmann et al., 2018*). With these forms of analysis, avoiding missing data in certain periods provides more evenly distributed statistical power across time points (even with ”guessing” of missing data through interpolation). Accordingly, *eyeris* uses the interpolate() function and takes the simplest approach of linear interpolation due to the relatively slow nature of pupil size fluctuations. However, users are strongly encouraged to adopt cutoff thresholds to exclude periods or trials when more than a certain percentage (such as 50%) of raw data is missing.

It is evident that the interpolation steps should occur after missing-data detection, but it is less clear if, when, and how smoothing of typically noisy pupillometry data should take place. Smoothing of pupil sizes over time is commonly employed in the literature (*Steinhauer et al., 2022*) both to provide cleaner time trace visualizations and to reduce sporadic temporal dynamics that may be unrelated to underlying neurocognitive processes of interest. Smoothing is justified for pupil sizes with relatively slow dynamics but should always occur after blink and artifact detection to avoid contaminating neighboring samples during smoothing. In addition, smoothing can be achieved in multiple different ways, such as with moving averages or low-pass filtering. *eyeris* uses the lpfilt() function to apply a low-pass Butterworth filter with passband starting at 4 Hz and stopband starting at 8 Hz. Note, the lpfilt() function here uses a minimum-order Butterworth filter, whose order is obtained with the buttord() function from the *gsignal* package (*van Boxtel, 2021*) to meet the defined passband and stopband cutoff frequencies (i.e., wp, ws) and ripple/attenuation requirements (i.e., rp, rs). The filter created by the *gsignal::butter()* function is returned in the second-order sections format. Users can easily modify the filter design parameters to achieve different smoothing results, but should always inspect the frequency response of the filter to ensure sufficient signal attenuation at higher frequencies and minimal ripples in the passband. The frequency response can be visualized by passing the plot_freqz = TRUE parameter to lpfilt().

Pupil sizes can fluctuate at a slow pace over the course of an entire recording; for example, the commonly reported decreasing trend during a cognitive task (a.k.a., the ”time-on-task effect”; *Beatty, 1982; Hopstaken et al., 2015; Martin et al., 2022; Massar et al., 2016; Steinhauer et al., 2022; Unsworth and Robison, 2016; Zhao et al., 2019*). Depending on the scientific questions at hand, removal of this time-on-task effect may be desirable. *eyeris* provides the detrend() function to remove a linear trend over the entire recording. However, it should be noted that pupillometry analysis often employs baseline correction in a trial-by-trial design; as such, detrending might be unnecessary and could also increase the risk of introducing artifacts in the data if an extreme trend gets extracted (although unlikely as these artifacts were likely removed in the preceding steps). Moreover, high-pass filtering is discouraged as a way to remove slow trends in pupillometry data because designing a filter with a sufficiently sharp cutoff at the lower frequency end is difficult; failing to implement a proper filter design could drastically change the waveforms of pupil sizes that contain frequency content at

approximately 1 Hz. Lastly, slow trends in pupil sizes might be physiologically meaningful, as tonic pupil sizes may reflect tonic arousal levels (*AstonJones and Cohen, 2005b; Groot et al., 2021; Joshi and Gold, 2020; Madore et al., 2020; Madore and Wagner, 2022; Murphy et al., 2011; Schwartz et al., 2025; Smallwood et al., 2012; Tsukahara et al., 2016; Unsworth and Robison, 2018*). Depending on the target research question(s), such trends may need to be preserved for downstream analysis. Therefore, *eyeris* defaults to not performing detrending during preprocessing.

Finally, there has been an ongoing debate regarding z-scoring normalization of pupillometry data. Some have argued against doing so outside of MRI contexts (*Steinhauer et al., 2022*). Normalizing pupil sizes by the mean and variance can introduce additional challenges in interpreting derived measures. Yet, it can also be argued that in the absence of normalization, pupil data across individuals cannot be easily pooled together for statistical inference, even if the non-normalized pupil sizes faithfully reflect individuals' physiological responses. To illustrate, older adults and young adults may have differences in physical constraints that impact how much the pupil can dilate during cognitive performance. Not accounting for such differences may lead to the conclusion that older adults have weaker pupillary responses to certain stimuli (such as a loud tone). However, if different ranges of pupil fluctuations across individuals are accounted for by using z-scored pupil sizes, older adults may show comparable responses to such stimuli relative to their normal ranges of pupillary responses (*He et al., 2020*). Therefore, while older adults may show attenuated changes in pupil size due to physiological changes in muscles controlling pupil dilation, their cognitive arousal response to tone stimuli may not be preferentially reduced to suggest a selective deficit. Similar examples can be found even within a homogeneous group, since differences across individuals and across blocks/sessions are often present in each recording (e.g., *Madore et al., 2020*). Averaging non-normalized pupillary responses produces a metric biased by absolute pupil size, potentially confounding the cognitive (pupil) signal of interest with physiological variability.

Hence, z-scoring pupillometry data allows one to focus on the relative dynamics of fluctuations in pupil sizes rather than the first-level characteristics such as mean pupil sizes or absolute changes. It should be emphasized that z-scoring automatically makes the assumption that the variances accountable by changes in the absolute value or range of pupil sizes are irrelevant to target scientific questions. In many modern applications of pupillometry within psychological sciences, this is not an unreasonable assumption (i.e., *Bradshaw, 1969; Gilzenrat et al., 2010; He et al., 2020; Peysakhovich et al., 2017; Van Gerven et al., 2004*). Therefore, *eyeris* defaults to perform z-scoring normalization with the zscore() function. One can easily turn off z-scoring during preprocessing, and we strongly encourage users to carefully consider how the decision to apply or forgo normalization of pupil sizes will impact downstream interpretations of their data, particularly when absolute pupil size or cross-session and cross-individual differences are relevant to the scientific question at hand.

Box 5. Tips & Tricks

Consider preprocessing your data with *eyeris::glassbox()* rather than manually calling individual steps.

Although we encourage most users to use the *glassbox()* function for simplicity and adherence to signal processing best practices, advanced users can create custom preprocessing steps that seamlessly integrate into the workflow (and thus will behave like native package functions without the need for hacky integrations). Therefore, it is imperative that the custom pipeline steps conform to the *eyeris* protocol to ensure maximum compatibility and reproducibility with downstream functions. We provide a thorough vignette that walks users through the process of creating custom extensions in the package documentation (*https://eyeris.shawnschwartz.com/articles/custom-extensions.html*) and in ***Section 3.1.6***.

3.1.2 A modular signal processing workflow with carefully chosen parameters ensures efficient and robust pupillometry preprocessing.

One key strength of the *eyeris* framework is its modularity. Each preprocessing function within *eyeris* is a wrapper around a core pipeline operation that automatically tracks, versions, and stores preprocessing states for any pupillometry dataset being processed in *eyeris*. This enables a ”plug-and-play” experience when using *eyeris*: users have the option to use our recommended glassbox() method, as well as the ability to construct novel extensions and custom pipelines. Critically, this modular architecture promotes both adaptability and extensibility, facilitating the streamlined integration of custom preprocessing methods directly into the default workflow (***Boxes 1,5,6,7; Appendix***).

We recommend that users start with the glassbox() function to reduce the probability of inadvertently introducing sub-optimal computations or errors that could arise when reimplementing the individual steps underlying the preprocessing pipeline – both within and between projects. The glassbox() function allows users to make flexible adjustments – such as disabling a particular step and/or changing the default settings for a step – by passing in parameter arguments (see ***Box 6***). A strength of the built-in glassbox() function compared to the many extant pupillometry packages is the framework it provides for rapidly preprocessing pupillometry data without requiring users to rely on copy-pasting complex step-by-step functions employed on a toy dataset from a vignette (which may not necessarily be optimized for the variety of forms that real-world pupillometry data come in). Furthermore, even when the instructions and functions supplied in a vignette of a package work ”out-of-the-box” after pasting them into a custom script, it is still the user’s responsibility to ensure that every step was not only run, but done so in the correct order (with respect to the best practices of signal processing). Failing to do so could lead to unexpected and possibly detrimental results for the conclusions researchers will ultimately draw from the data. Therefore, we aim to facilitate this process to enable users to more easily engage with preprocessing steps by using a package, rather than moving away from them. With the *eyeris* glassbox() function, users can take a pupillometry recording file directly

from a participant session, and in a short amount of time can go from raw to fully preprocessed data with detailed diagnostics for quality control in approximately three lines of code (***Box 1***,***5***; ***Figure 2***).

However, one critique of encapsulating a series of preprocessing steps into a single wrapper function – like glassbox() – is the regression towards a more black-box-like implementation. *eyeris* circumvents this pitfall in two key ways. First, the glassbox() function has two modes: interactive and headless. The interactive mode allows users to evaluate the consequences of the parameters used within each step on their data in real time, thereby facilitating an easy-to-use workflow for parameter optimization on any given dataset. In essence, this interactive process takes each recommended step and forces the user to evaluate both pre- and post-processing plots of time series for all steps in the default pipeline, with interjecting prompts to continue or modify parameters before moving forward. This stepwise approach to optimizing signal preprocessing parameters that yield satisfactorily de-noised pupil time series data reduces the opacity of the wrapper function to a transparent and engaging experience. The headless mode permits users to run the pipeline as-is on a dataset of any size – be it on their computer locally or on a cloud computing cluster – to preprocess large-scale pupillometry datasets using our recommended default parameters without manual intervention.

Box 6. Overriding default parameters in the *eyeris::glassbox()*

```
output <- glassbox(
  rawdata,
  confirm = FALSE,
  deblink = list(extend = 40),
  detransient = FALSE,
  lpfilt = list(plot_freqz = FALSE)
)
#> [  OK  ] - Running eyeris::load_asc()
#> [  OK  ] - Running eyeris::deblink()
#> [ SKIP ] - Skipping eyeris::detransient()
#> [  OK  ] - Running eyeris::interpolate()
#> [  OK  ] - Running eyeris::lpfilt()
#> [ SKIP ] - Skipping eyeris::detrend()
#> [  OK  ] - Running eyeris::zscore()
```

Second, *eyeris* stores the history of preprocessing steps and parameters for every action made on a raw dataset so that any derived data from the pipeline can be easily reproduced without additional steps needed from the researcher. Specifically, all of these data are saved to interactive HTML files (***Figure 2***), which contain numerous diagnostic plots for every step of the pipeline, automatically parsed tracker metadata, and a snapshot of the actions performed at each step of the workflow (see ***Section 3.1.4***). Thus, using *eyeris* to deploy a large-scale preprocessing workflow on a headless computing architecture allows researchers to manually scrub their data sets for each run from every subject at every step of preprocessing without additional custom scripting. Critically, these interactive reports provided by *eyeris* go above and beyond within-study data quality control by automatically

capturing and recording data and signal processing provenance in a standardized schema (*MacKenzie-Graham et al., 2008*) – an essential feature for analytic reproducibility, lacking in many existing pupillometry preprocessing tools.

3.1.3 User-friendly pattern matching and metadata parsing make extracting trial level epochs both accurate and effortless.

Extant solutions for preprocessing pupil data lack intuitive workflows to extract trial-level epochs. For EEG researchers used to functions like pop_epoch in the MATLAB-based *EEGLAB* software (*Delorme and Makeig, 2004*), the lack of a comparable analogous function can lead to bottlenecks in a researcher's pupillometry preprocessing workflow. Naturally, it made sense to take inspiration from *EEGLAB* when designing the epoch() function within *eyeris*. In particular, extracting trial-level data epochs with *eyeris* is flexible and offers advanced tooling for the extraction of metadata from event message strings. To illustrate, a researcher may have event message tags that look something like: PROBE_START_32_PIANO, indicating the timestamp for the stimulus probe onset on trial 32 where the stimulus was the piano. While not all event message strings contain as much information as this contrived example (e.g., standard EyeLink trial markers *TRIALID 1* often provide helpful trial-number metadata), the additional sources of information in the former example can be leveraged by the researcher as an additional sanity check to ensure accurate linking of pupil data with behavioral and/or neuroimaging data prior to analysis. Unsurprisingly, with more complex event message strings comes more burden on the researcher to devise an accurate and reliable solution to handle their specific event message configuration, likely leading to inefficient one-off solutions for particular datasets. *eyeris* meets this common demand by providing a carefully crafted regular expression implementation designed to handle a wide variety of generic cases (***Box 3***).

Conceptually, there are numerous ways that a researcher might want to extract their trial epochs for downstream analysis. In the simplest scenario, the user may be interested in testing a hypothesis about the tonic pupil size in the 1-second period just before the onset of the stimulus. In this case, the user would need to provide the string pattern and fixed start/stop duration of the epoch they would like to extract. For example, c(-1,0) for the 1-second just before the onset of the trial specifies a start/stop time array using the matched event message strings as the reference point (i.e., time=0).

3.1.4 Interactive visual reports and a BIDS-like data scheme enable intuitive quality control and transparent data reporting.

To date, most tools for preparing pupillometry data lack suitable features for data quality control up to the individual trial level. While some packages have integrated helpful plotting functions (e.g., *Forbes, 2020; Hershman et al., 2019; Mittner, 2020; Tsukahara, 2020*), these approaches only get the user part of the way there, as the onus falls on the user to aggregate such plots in a meaningful way for quality control. Moreover, this significant hurdle could be an arduous task even for researchers with

programming experience. We strongly recommend that researchers carefully inspect the data at each stage of the preprocessing pipeline before aggregating trials and/or running inferential statistics, as is common in other human neuroscience data modalities such as EEG with the *EEGLAB* interactive viewer (*Delorme and Makeig, 2004*), M/EEG with *MNE-Python* (*Gramfort et al., 2013*), fMRI with *fMRIPrep* (*Esteban et al., 2019*), and structural MRI with *MRIQC* (*Esteban et al., 2017; Provins et al., 2023*) visual reports.

Given the lack of generic plug-and-play quality control solutions to date – in addition to the amount of effort required to implement such a solution – we suspect that many published pupillometry studies may not have raw data readily inspected end-to-end before and after preprocessing and running statistics. We therefore provide users with visualization tools of the pupil time series before and after each preprocessing operation for both an entire block/task and every trial-level epoch (i.e., as delineated by event messages/epochs). With this key feature, we hope to lower the barrier required for researchers to visually inspect entire recording sessions for anomalies in an efficient manner, as is necessary to ensure reproducible research results.

For the above reasons, we made visual reporting the default behavior in *eyeris* (***Figure 2***). Similar to the goals put forth by the team behind *fMRIPrep*, we anticipate that automatically generating diagnostic reports per participant will (*i*) maximize shareability and transparency between researchers, (*ii*) make data surfing less cumbersome and more efficient and hence increase the probability that thorough data surfing occurs before submitting data to statistical tests, (*iii*) increase statistical power of future pupillometry investigations by enabling researchers to purge poor data at the trial level without excluding entire runs or participants altogether, and (*iv*) contribute to best practices in open and reproducible science by recording all preprocessing decisions and settings for any given dataset that makes its way through *eyeris* by default (*Esteban et al., 2020, 2019; Poldrack et al., 2019*).

Importantly, all derived data and reports from the foregoing preprocessing workflows are automatically sorted into a BIDS-like format (*Gorgolewski et al., 2016*), facilitating separation of source (i.e., original or raw) data from derived (i.e., preprocessed, trimmed, epoched, etc.) data, in a directory structure that is intuitive to navigate, automatically organized, and therefore essentially ready to share on an appropriate data sharing outlet/repository like OSF, dryad, and/or GitHub (see ***Box 4***).

3.1.5 Integrated database storage enables scalable and efficient management of large-scale pupillometry datasets.

As pupillometry research increasingly involves large sample sizes, multi-session designs, and high-throughput cloud computing deployments, traditional CSV-based data management workflows can become unwieldy. For studies with hundreds or thousands of participants – each generating multiple recording files, preprocessing derivatives, epoch extractions, and diagnostic metadata – the resulting proliferation of individual CSV files poses challenges for data organization, storage efficiency, and analytic tractability. Moreover, syncing and transferring thousands of small files across cloud

storage systems incurs significant I/O overhead, and loading entire datasets into memory for analysis is often impractical.

To address these scalability challenges, *eyeris* provides an integrated database storage backend powered by DuckDB (*Raasveldt & Mühleisen, 2019*), a high-performance embedded analytical database engine. When database storage is enabled within the *bidsify()* function (by setting *db_enabled* = *TRUE*), all preprocessed data – including time series, events, blinks, epochs, epoch summaries, and confounds – are written to a single, self-contained database file (.*eyerisdb*). This centralized storage approach offers several key advantages over flat-file workflows. First, queries execute at the storage level, allowing researchers to filter, aggregate, and join data across subjects without loading entire datasets into memory. Second, a single database file dramatically reduces I/O costs on cloud platforms compared to syncing thousands of individual files. Third, built-in schema validation maintains data integrity and automatically tracks metadata such as creation timestamps, data types, and relationships between tables. Fourth, researchers can leverage SQL for complex cross-subject filtering and targeted extraction of data subsets, while the *eyeris_db_collect()* wrapper function provides a user-friendly interface that requires no SQL knowledge.

Critically, the database workflow is designed to complement – not replace – the existing BIDS-like file structure described above. Users may choose to generate both CSV files and a database simultaneously, maintaining human-readable outputs alongside the database for transparency, or opt for a database-only workflow to minimize storage and I/O costs in cloud computing environments (e.g., when deploying *eyeris* on high-performance computing clusters). This flexibility ensures that *eyeris* can scale from small desktop analyses involving a handful of participants to large-scale, multi-site pupillometry studies with hundreds or thousands of recordings, without requiring researchers to fundamentally change their preprocessing workflow.

For collaborative research and open-science workflows, *eyeris* also provides tools for distributing large databases in portable formats. The *eyeris_db_to_chunked_files()* function enables chunked export of large databases to CSV or Parquet files, processing data incrementally to avoid memory constraints and automatically splitting files that exceed a configurable size threshold. The Parquet format, in particular, offers substantial advantages for archival and sharing: 50–80% file size reduction compared to CSV, faster read performance, and preserved numeric precision. Together with *eyeris_db_split_for_sharing()* and *eyeris_db_reconstruct_from_chunks()*, these tools enable seamless data distribution and reconstruction, further supporting the FAIR principles that underpin *eyeris*.

3.1.6 Advanced features for developers.

We provide developers the flexibility to extend upon *eyeris*' default offerings by writing their own custom pipeline operations that can integrate seamlessly with *eyeris* primitives (see ***Box 7***).

Box 7. Deconstructing *glassbox()*: Using the flexible *eyeris* primitives

```
system.file("extdata", "memory.asc", package = "eyeris") |>
  load_asc(block = "auto") |>
  deblink(extend = 50) |>
  detransient(n = 16) |>
  interpolate() |>
  lpfilt(wp = 4, ws = 8, rp = 1, rs = 35, plot_freqz = TRUE) |>
  # eyeris::detrend() |>  # optional (read docs before enabling)
  zscore()
```

Under the hood, each preprocessing function in *eyeris* is a wrapper around a core operation that gets tracked, versioned, and stored using the *pipeline_handler()* function. The *pipeline_handler()* enables flexible integration of custom data processing functions into the *eyeris* pipeline. It accepts an *eyeris* class object, the name of the operation function to apply, a character string suffix for the new column name, and any additional custom parameters. The function returns an updated *eyeris* object with the new column added to the *timeseries* data frame and the *latest* pointer updated to reflect the full history of preprocessing steps (e.g., *pupil_raw* → *pupil_raw_deblink* → *pupil_raw_deblink_detransient* → *pupil_...*). Following the *eyeris* protocol ensures that all operations follow a predictable structure and that new pupil data columns based on previous operations in the chain are dynamically constructed within the core time series data frame. Custom pipeline steps must conform to this protocol for maximum compatibility with the downstream functions we provide.

To demonstrate the flexibility of *eyeris* for developers, we walk through building a custom extension below. Please note that in the example that follows, we are *not* recommending that one performs this particular statistical transformation on their pupil data; rather, this is purely for demonstration purposes. Suppose you want to write a new *eyeris* extension function called *winsorize()* to apply winsorization[6] to extreme pupil values. First, you write the core operation function, which should accept a data frame *x*, a string *prev_op* (the name of the previous pupil column), and any custom parameters. Second, you create a wrapper using the *pipeline_handler()* function, which enables your function to automatically: (a) track your function within the *eyeris* list object's *params* field, (b) append a new column to each block within the *timeseries* list of data frames, and (c) update the object's *latest* pointer. To ensure compatibility with *eyeris*, we impose the following requirements for any custom extension functions: (1) the first argument is always the *eyeris* class object; (2) the second is the name of the core operation function; (3) the third is the internal label you want *eyeris* to refer to (i.e., the one for the column name, plots, etc.); and (4) the fourth position onward are the custom parameters you defined in the core operation function. See ***Appendix*** for all source code for this demo. More information about customizing *eyeris* can be found in the online package documentation[7].

[6] https://en.wikipedia.org/wiki/Winsorizing
[7] https://eyeris.shawnschwartz.com/articles/custom-extensions.html

4 Conclusions

The bells and whistles detailed throughout this manuscript and the documentation website[8] for the *eyeris* R package aim to standardize the way pupillometry data are preprocessed and used in downstream analysis. To this end, we designed the architecture underlying *eyeris* to support the building of a robust suite of pupillometry tools analogous to the current offerings for fMRI analysts with *fMRIPrep* (*Esteban et al., 2020, 2019*). At its core, *eyeris* enables standardized and flexible preprocessing of the pupil assay.

Moreover, *eyeris* was designed with FAIR[9] principles in mind. To fulfill this promise, *eyeris* uses a built-in protocol to carry out and manage the record keeping of the entire preprocessing pipeline on a raw dataset (see ***Figure 1***). This protocol is more or less a ”grammar”, where each operation is a ”verb” that can be chained one after the other (see ***Box 7***, which draws inspiration from familiar tidyverse ”pipe” workflows; *Wickham et al., 2019*). This approach allows us to take advantage of syntactic sugar such that one could take a raw pupillometry dataset and within seconds obtain (*i*) a step-by-step record of all preprocessing steps (and parameters used), (*ii*) how each step sequentially impacted the raw signal, (*iii*) default diagnostic reports that organize relevant preprocessing metadata for any particular dataset passed through *eyeris*, and (*iv*) an intuitive gallery viewing functionality for efficiently scrolling through hundreds-to-thousands of raw data epochs for quality control prior to running inferential statistics (***Figure 2***).

Additionally, *eyeris* provides integrated database storage powered by DuckDB, enabling researchers to manage and query preprocessed pupillometry data from large-scale studies – involving hundreds or thousands of participants – within a single, self-contained database file. This database backend, combined with chunked export tools for portable data sharing in CSV and Parquet formats, ensures that *eyeris* can scale from small desktop analyses to cloud-deployed, multi-site studies without compromising on the transparency, reproducibility, or accessibility of the preprocessing workflow.

To this end, our hope is that enabling access to intuitive, interactive visual inspection tools to efficiently assess trial/epoch-level data quality – combined with scalable data management infrastructure – at scale, will support researchers in their efforts to perform reproducible and rigorous pupillometry research in a more accessible manner than is afforded by extant offerings.

---

[8] https://eyeris.shawnschwartz.com

[9] Findability, Accessibility, Interoperability, and Reusability (see *Hahn et al., 2024; Lazar, 2023; Wilkinson et al., 2016*).

Acknowledgments

Development of this software library was supported by an agility grant from the Stanford University Wu Tsai Human Performance Alliance (to Prof. Anthony D Wagner & Shawn T Schwartz); the Stanford University Wu Tsai Neurosciences Institute Center for Mind, Brain, Computation and Technology (to Shawn T Schwartz & Alice M Xue); a Stanford University Ric Weiland Graduate Fellowship in the Humanities & Sciences (to Shawn T Schwartz); the McKnight Brain Research Foundation Clinical Translational Research Scholarship in Cognitive Aging and Age-Related Memory Loss; the American Brain Foundation; the American Academy of Neurology (to Haopei Yang); National Science Foundation Graduate Research Fellowship DGE-2146755 (to Alice M Xue); a Stanford Interdisciplinary Graduate Fellowship (to Alice M Xue); and National Institutes of Health Grants R01-AG065255 (to Prof. Anthony D Wagner) and R01-AG079345 (to Prof. Patrick L Purdon). The funders had no role in study design, data collection, and interpretation, or the decision to submit the work for publication. Furthermore, we are grateful to Ryan Yan for providing helpful comments on this manuscript. We are also grateful to Anthony Wagner for supporting the development of this software. We are thankful to Gustavo Santiago-Reyes for debugging and contributing to an earlier version of the source code, as well as to Natalie Wang for beta-testing the software during its early stages of development. We additionally thank Brian Knutson for helpful theoretical and practical discussions about pupillometry and arousal mechanisms. Lastly, we extend our gratitude to the past and present members of the Stanford Memory Lab[10] who have influenced the work discussed herein.

Author Contributions

Shawn T Schwartz, Conceptualization, Formal Analysis, Methodology, Project Administration, Resources, Software, Supervision, Validation, Visualization, Writing - Original Draft Preparation, Writing - Review & Editing; **Haopei Yang**, Data Curation, Resources, Validation, Writing - Review & Editing; **Alice M Xue**, Data Curation, Resources, Validation, Writing - Review & Editing; **Mingjian He**, Formal Analysis, Methodology, Resources, Supervision, Validation, Writing - Review & Editing.

Code Availability

Please use the *issues* tab on GitHub to note bugs, comments, suggestions, and/or feedback.

- CRAN: *https://cran.r-project.org/package=eyeris*
- GitHub repository: *https://github.com/shawntz/eyeris*
- GitHub issues: *https://github.com/shawntz/eyeris/issues*
- Package documentation: *https://eyeris.shawnschwartz.com*

[10] *https://memorylab.stanford.edu*

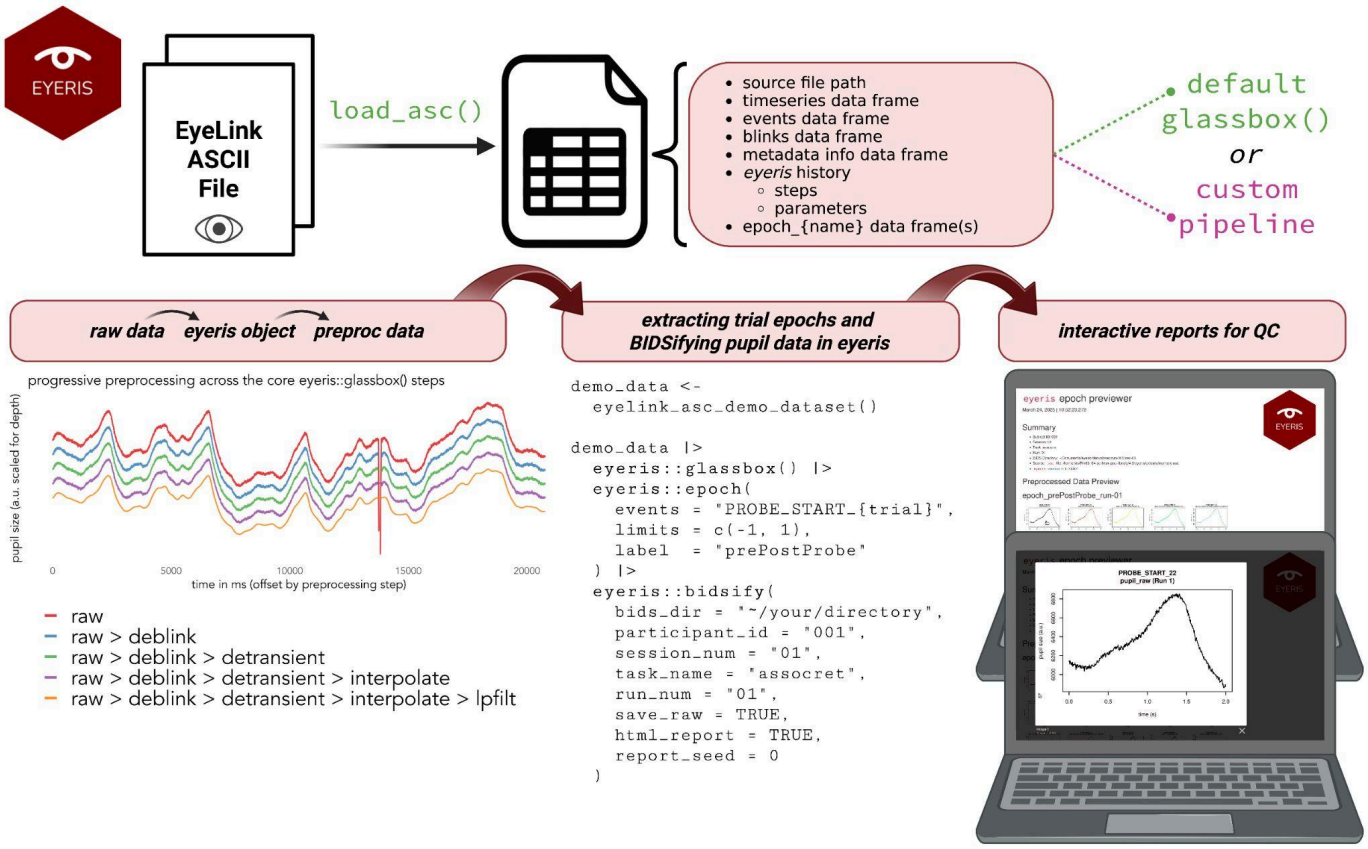

Figure 1. A sample workflow for the *eyeris* R package. The example code shows an end-to-end preprocessing pipeline, starting at raw data from the tracker to epoched and derived files ready for visualization and analysis. Here, we draw particular attention to the flexibility *eyeris* offers for segmenting (i.e., epoching) trials from the continuous pupil time series. In this demonstration, the *epoch() function* will find all event messages from the recording session that include the "PROBE_START_*" wildcard expression and extract 1-second prior to as well as 1-second following the onset timestamp from the matched message string (and will name the resulting tidy data frame of extracted pupil epochs: "prePostProbe", which will be saved to the resulting *eyeris* S3 class "list" object — *not shown here*). Note, the "{trial}" syntax is a feature that enables a user to tell *eyeris* to not only treat that portion of the string as a wildcard expression, *but also* to take whatever substring matched to that position of the string and place its content in a new column with the header title: *trial*. This feature was designed with the intention to streamline the extraction of metadata embedded within tracker event messages into the same tidy data frame as any corresponding pupil data extracted from these events. The *epoch()* function is quite flexible and can handle a wide range of event message-string configurations; an in depth walkthrough can be found in the package documentation. Created in BioRender. *Schwartz (2025b)* *https://BioRender.com/vkahl3u*.

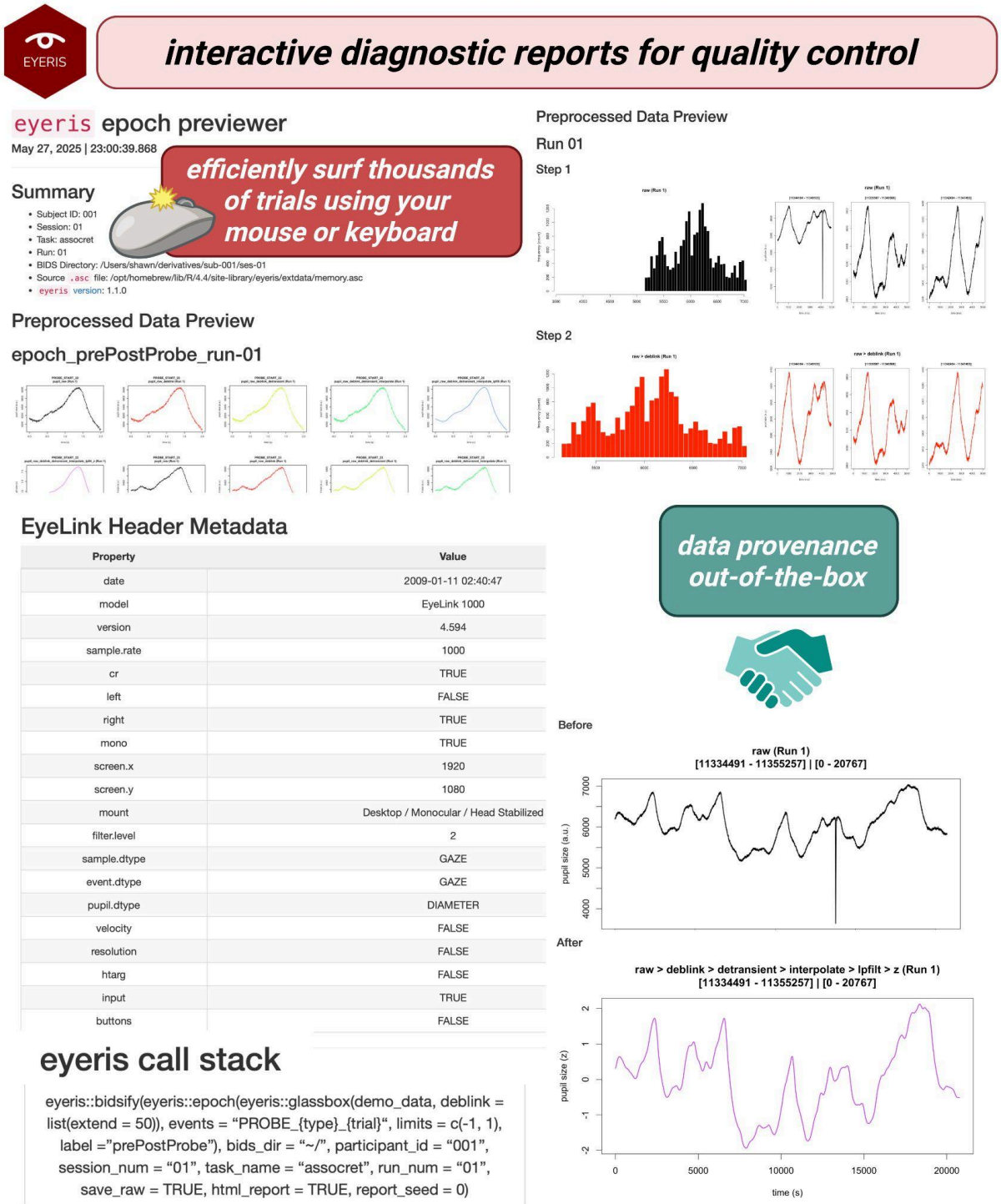

Figure 2. Overview of data and signal processing provenance provided out-of-the-box. *eyeris* provides automatically generated summary reports with diagnostic plots, recording metadata, and all preprocessing operations/parameters — and the resulting effects of each operation on the pupil time series — for every file/run for every participant. *eyeris* reports facilitate quality control by enabling surfing through thousands of trials worth of pupil data efficiently and interactively. Furthermore, these reports provide data provenance out-of-the-box and can be shared alongside any publication to support best practices in open science and reproducibility. Created in BioRender. *Schwartz (2025c)* https://BioRender.com/j2h2hzs.

References

Alamia A, VanRullen R, Pasqualotto E, Mouraux A, Zenon A. Pupil-Linked Arousal Responds to Unconscious Surprisal. The Journal of Neuroscience. 2019; 39(27):5369–5376. https://doi.org/10.1523/jneurosci.3010-18.2019.

Arachchige CN, Prendergast LA. Confidence intervals for median absolute deviations. Communications in Statistics-Simulation and Computation. 2024; p. 1–10. https://doi.org/10.1080/03610918.2024.2376198.

Aston-Jones G, Cohen JD. An Integrative Theory Of Locus Coeruleus-Norepinephrine Function: Adaptive Gain and Optimal Performance. Annual Review of Neuroscience. 2005; 28(1):403–450. https://doi.org/10.1146/annurev.neuro.28.061604.135709.

Aston-Jones G, Cohen JD. An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. Annu Rev Neurosci. 2005; 28(1):403–450. https://doi.org/10.1146/annurev.neuro.28.061604.135709.

Beatty J. Phasic Not Tonic Pupillary Responses Vary With Auditory Vigilance Performance. Psychophysiology. 1982; 19(2):167–172. https://doi.org/10.1111/j.1469-8986.1982.tb02540.x.

Betts MJ, Kirilina E, Otaduy MC, Ivanov D, Acosta-Cabronero J, Callaghan MF, Lambert C, Cardenas-Blanco A, Pine K, Passamonti L, et al. Locus coeruleus imaging as a biomarker for noradrenergic dysfunction in neurodegenerative diseases. Brain. 2019; 142(9):2558–2571. https://doi.org/10.1093/brain/awz193.

Bigdely-Shamlo N, Mullen T, Kothe C, Su KM, Robbins KA. The PREP pipeline: standardized preprocessing for large-scale EEG analysis. Frontiers in neuroinformatics. 2015; 9(16):1–20. https://doi.org/10.3389/fninf.2015.00016.

van Boxtel G, gsignal: Signal processing; 2021. https://doi.org/10.32614/CRAN.package.gsignal.

Braak H, Del Tredici K. Alzheimer's pathogenesis: is there neuron-to-neuron propagation? Acta neuropathologica. 2011; 121:589–595. https://doi.org/10.1007/s00401-011-0825-z.

Bradley MM, Miccoli L, Escrig MA, Lang PJ. The pupil as a measure of emotional arousal and autonomic activation. Psychophysiology. 2008; 45(4):602–607. https://doi.org/10.1111/j.1469-8986.2008.00654.x.

Bradshaw JL. Background light intensity and the pupillary response in a reaction time task. Psychonomic Science. 1969; 14(6):271–272. https://doi.org/10.3758/BF03329118.

Chen A, Wong C, Sharif B, Peruma A. Exploring Code Comprehension in Scientific Programming: Preliminary Insights from Research Scientists. arXiv preprint arXiv:250110037. 2025; https://doi.org/10.48550/arXiv.2501.10037.

Clewett D, Gasser C, Davachi L. Pupil-linked arousal signals track the temporal organization of events in memory. Nature Communications. 2020; 11(1):4007. https://doi.org/10.1038/s41467-020-17851-9.

Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of neuroscience methods. 2004; 134(1):9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009.

Einhäuser W, Stout J, Koch C, Carter O. Pupil dilation reflects perceptual selection and predicts subsequent stability in perceptual rivalry. Proceedings of the National Academy of Sciences. 2008; 105(5):1704–1709. https://doi.org/10.1073/pnas.0707727105.

El Haj M. Pupillometry as tool to assess cognitive and affective processing in aging. Brain Disorders. 2024; p. 100129. https://doi.org/10.1016/j.dscb.2024.100129.

Eldar E, Cohen JD, Niv Y. The effects of neural gain on attention and learning. Nature Neuroscience. 2013; 16(8):1146–1153. https://doi.org/10.1038/nn.3428.

Esteban O, Birman D, Schaer M, Koyejo OO, Poldrack RA, Gorgolewski KJ. MRIQC: Advancing the automatic prediction of image quality in MRI from unseen sites. PloS one. 2017; 12(9):e0184661. https://doi.org/10.1371/journal.pone.0184661.

Esteban O, Ciric R, Finc K, Blair RW, Markiewicz CJ, Moodie CA, Kent JD, Goncalves M, DuPre E, Gomez DEP, Ye Z, Salo T, Valabregue R, Amlien IK, Liem F, Jacoby N, Stojić H, Cieslak M, Urchs S, Halchenko YO, et al. Analysis of task-based functional MRI data preprocessed with fMRIPrep. Nature Protocols. 2020; 15(7):2186–2202. https://doi.org/10.1038/s41596-020-0327-3.

Esteban O, Markiewicz CJ, Blair RW, Moodie CA, Isik AI, Erramuzpe A, Kent JD, Goncalves M, DuPre E, Snyder M, Oya H, Ghosh SS, Wright J, Durnez J, Poldrack RA, Gorgolewski KJ. fMRIPrep: a robust preprocessing pipeline for functional MRI. Nature Methods. 2019; 16(1):111–116. https://doi.org/10.1038/s41592-018-0235-4.

Fink L, Simola J, Tavano A, Lange E, Wallot S, Laeng B. From pre-processing to advanced dynamic modeling of pupil data. Behavior Research Methods. 2024; 56(3):1376–1412. https://doi.org/10.3758/s13428-023-02098-1.

Forbes SH. PupillometryR: An R package for preparing and analysing pupillometry data. Journal of Open Source Software. 2020; 5(50):2285. https://doi.org/10.21105/joss.02285.

Frank MC, Braginsky M, Cachia J, Coles NA, Hardwicke TE, Hawkins RD, Mathur MB, Williams R. Replication. In: Experimentology: An Open Science Approach to Experimental Psychology Methods Stanford University; 2025.https://doi.org/10.25936/3JP6-5M50, also published by MIT Press, ISBN 978-0-262-55256-1.

Gee JWd, Knapen T, Donner TH. Decision-related pupil dilation reflects upcoming choice and individual bias. Proceedings of the National Academy of Sciences. 2014; 111(5):E618–E625. https://doi.org/10.1073/pnas.1317557111.

Geller J, Winn MB, Mahr T, Mirman D. GazeR: A Package for Processing Gaze Position and Pupil Size Data. Behavior Research Methods. 2020; 52(5):2232–2255. https://doi.org/10.3758/s13428-020-01374-8.

Gelman A, Loken E. The garden of forking paths: Why multiple comparisons can be a problem, even when there is no “fishing expedition” or “p-hacking” and the research hypothesis was posited ahead of time. Department of Statistics, Columbia University. 2013; 348(1-17):3. https://osf.io/n3axs.

Gilzenrat MS, Nieuwenhuis S, Jepma M, Cohen JD. Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. Cognitive, Affective, & Behavioral Neuroscience. 2010; 10(2):252–269. https://doi.org/10.3758/CABN.10.2.252.

Gorgolewski KJ, Auer T, Calhoun VD, Craddock RC, Das S, Duff EP, Flandin G, Ghosh SS, Glatard T, Halchenko YO, Handwerker DA, Hanke M, Keator D, Li X, Michael Z, Maumet C, Nichols BN, Nichols TE, Pellman J, Poline JB, et al. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. Scientific Data. 2016; 3(1):160044. https://doi.org/10.1038/sdata.2016.44.

Gorgolewski KJ, Poldrack RA. A practical guide for improving transparency and reproducibility in neuroimaging research. PLoS biology. 2016; 14(7):e1002506. https://doi.org/10.1371/journal.pbio.1002506.

Gramfort A, Luessi M, Larson E, Engemann DA, Strohmeier D, Brodbeck C, Goj R, Jas M, Brooks T, Parkkonen L, et al. MEG and EEG data analysis with MNE-Python. Frontiers in Neuroinformatics. 2013; 7:267. https://doi.org/10.3389/fnins.2013.00267.

Groot JM, Boayue NM, Csifcsák G, Boekel W, Huster R, Forstmann BU, Mittner M. Probing the neural signature of mind wandering with simultaneous fMRI-EEG and pupillometry. NeuroImage. 2021; 224:117412. https://doi.org/10.1016/j.neuroimage.2020.117412.

Haab O. Vortrag gesellschaft der ärzte in zürich am 21. November 1885. Korrespondenzblatt für Schweizer Ärzte. 1886; 16:153.

Hahn S, Allgaier N, Garavan H. Methods for Analyzing Large Neuroimaging Datasets. Neuromethods. 2024; p. 255–287. https://doi.org/10.1007/978-1-0716-4260-3_12.

Hampel FR. The influence curve and its role in robust estimation. Journal of the american statistical association. 1974; 69(346):383–393.

Hampel FR. Contributions to the theory of robust estimation. University of California, Berkeley; 1968.

Hardwicke TE, Bohn M, MacDonald K, Hembacher E, Nuijten MB, Peloquin BN, DeMayo BE, Long B, Yoon EJ, Frank MC. Analytic reproducibility in articles receiving open data badges at the journal Psychological Science: An observational study. Royal Society open science. 2021; 8(1):201494. https://doi.org/10.1098/rsos.201494.

Hardwicke TE, Mathur MB, MacDonald K, Nilsonne G, Banks GC, Kidwell MC, Hofelich Mohr A, Clayton E, Yoon EJ, Henry Tessler M, et al. Data availability, reusability, and analytic reproducibility: Evaluating the impact of a mandatory open data policy at the journal Cognition. Royal Society open science. 2018; 5(8):180448. https://doi.org/10.1098/rsos.180448.

Hardwicke TE, Thibault RT, Kosie JE, Wallach JD, Kidwell MC, Ioannidis JP. Estimating the prevalence of transparency and reproducibility-related research practices in psychology (2014–2017). Perspectives on Psychological Science. 2022; 17(1):239–251. https://doi.org/10.1177/1745691620979806.

Hatton L. The T-experiments: errors in scientific software. Quality of Numerical Software: Assessment and Enhancement. 1997; p. 12–31. https://doi.org/10.1007/978-1-5041-2940-4_2.

Hatton L, Roberts A. How accurate is scientific software? IEEE Transactions on Software Engineering. 1994; 20(10):785–797. https://doi.org/10.1109/32.328993.

He M, Heindel WC, Nassar MR, Siefert EM, Festa EK. Age-related changes in the functional integrity of the phasic alerting system: a pupillometric investigation. Neurobiology of aging. 2020; 91:136–147. https://doi.org/10.1016/j.neurobiolaging.2020.02.025.

Hershman R, Henik A, Cohen N. CHAP: Open-source software for processing and analyzing pupillometry data. Behavior Research Methods. 2019; 51(3):1059–1074. https://doi.org/10.3758/s13428-018-01190-1.

Hershman R, Milshtein D, Henik A. Processing and Analyzing of Pupillometry Data. In: Modern Pupillometry, Cognition, Neuroscience, and Practical Applications Springer; 2024.p. 431–466. https://doi.org/10.1007/978-3-031-54896-3_15.

Holper S, Watson R, Yassi N. Tau as a biomarker of neurodegeneration. International Journal of Molecular Sciences. 2022; 23(13):7307. https://doi.org/10.3390/ijms23137307.

Hopstaken JF, van der Linden D, Bakker AB, Kompier MAJ. A multifaceted investigation of the link between mental fatigue and task disengagement. Psychophysiology. 2015; 52(3):305–315. https://doi.org/10.1111/psyp.12339.

Huang R, Clewett D. The Locus Coeruleus: Where Cognitive and Emotional Processing Meet the Eye. In: Modern Pupillometry, Cognition, Neuroscience, and Practical Applications Springer; 2024.p. 3–75. https://doi.org/10.1007/978-3-031-54896-3_1.

Ioannidis JP. Why most published research findings are false. PLoS medicine. 2005; 2(8):e124. https://doi.org/10.1371/journal.pmed.0020124.

Jepma M, Nieuwenhuis S. Pupil Diameter Predicts Changes in the Exploration–Exploitation Trade-off: Evidence for the Adaptive Gain Theory. Journal of Cognitive Neuroscience. 2011; 23(7):1587–1596. https://doi.org/10.1162/jocn.2010.21548.

Johansson R, Parnamets P, Bjernestedt A, Johansson M. Pupil dilation tracks the dynamics of mnemonic interference resolution. Scientific reports. 2018; 8(1):4826. https://doi.org/10.1038/s41598-018-23297-3.

Joshi S, Gold JI. Pupil size as a window on neural substrates of cognition. Trends in cognitive sciences. 2020; 24(6):466–480. https://doi.org/10.1016/j.tics.2020.03.005.

Joshi S, Li Y, Kalwani RM, Gold JI. Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. Neuron. 2016; 89(1):221–234. https://doi.org/10.1016/j.neuron.2015.11.028.

Kahneman D, Beatty J. Pupil Diameter and Load on Memory. Science. 1966; 154(3756):1583–1585. https://doi.org/10.1126/science.154.3756.1583.

Kawaguchi K, Clery S, Pourriahi P, Seillier L, Haefner RM, Nienborg H. Differentiating between Models of Perceptual Decision Making Using Pupil Size Inferred Confidence. The Journal of Neuroscience. 2018; 38(41):8874– 8888. https://doi.org/10.1523/jneurosci.0735-18.2018.

Keene PA, deBettencourt MT, Awh E, Vogel EK. Pupillometry signatures of sustained attention and working memory. Attention, Perception, & Psychophysics. 2022; 84(8):2472–2482. https://doi.org/10.3758/s13414-022-02557-5.

Kim J, Lee J, Jun SB, Sung JE. Pupillometry as a window to detect cognitive aging in the brain. Biomedical Engineering Letters. 2024; 14(1):91–101. https://doi.org/10.1007/s13534-023-00315-6.

Kinley I, Levy Y. PuPl: an open-source tool for processing pupillometry data. Behavior Research Methods. 2021; p. 1–24. https://doi.org/10.3758/s13428-021-01717-z.

Knapen T, Gee JWd, Brascamp J, Nuiten S, Hoppenbrouwers S, Theeuwes J. Cognitive and Ocular Factors Jointly Determine Pupil Responses under Equiluminance. PLoS ONE. 2016; 11(5):e0155574. https://doi.org/10.1371/journal.pone.0155574.

Kovacs M, Varga MA, Dianovics D, Poldrack RA, Aczel B. Opening the black box of article retractions: exploring the causes and consequences of data management errors. Royal Society Open Science. 2024; 11(12):240844. https://doi.org/10.1098/rsos.240844.

Kramer SE, Lorens A, Coninx F, Zekveld AA, Piotrowska A, Skarzynski H. Processing load during listening: The influence of task characteristics on the pupil response. Language and Cognitive Processes. 2013; 28(4):426– 442. https://doi.org/10.1080/01690965.2011.642267.

Kret ME, Sjak-Shie EE. Preprocessing pupil size data: Guidelines and code. Behavior Research Methods. 2019; 51(3):1336–1342. https://doi.org/10.3758/s13428-018-1075-y.

Kristjansson SD, Stern JA, Brown TB, Rohrbaugh JW. Detecting phasic lapses in alertness using pupillometric measures. Applied Ergonomics. 2009; 40(6):978–986. https://doi.org/10.1016/j.apergo.2009.04.007.

Laeng B, Mathôt S. Methodological Aspects of Pupillometry. In: Modern Pupillometry: Cognition, Neuroscience, and Practical Applications Springer, Cham; 2024.p. 375–400. https://doi.org/10.1007/978-3-031-54896-3_13.

Lazar NA. Functional neuroimaging in the era of Big Data and Open Science: A modern overview. Wiley Interdisciplinary Reviews: Computational Statistics. 2023; 15(5). https://doi.org/10.1002/wics.1609.

Leys C, Ley C, Klein O, Bernard P, Licata L. Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. Journal of experimental social psychology. 2013; 49(4):764-766. https://doi.org/10.1016/j.jesp.2013.03.013.

MacKenzie-Graham AJ, Van Horn JD, Woods RP, Crawford KL, Toga AW. Provenance in neuroimaging. Neuroimage. 2008; 42(1):178–195. https://doi.org/10.1016/j.neuroimage.2008.04.186.

Madore KP, Khazenzon AM, Backes CW, Jiang J, Uncapher MR, Norcia AM, Wagner AD. Memory failure predicted by attention lapsing and media multitasking. Nature. 2020; 587(7832):87–91. https://doi.org/10.1038/s41586-020-2870-z.

Madore KP, Wagner AD. Readiness to remember: predicting variability in episodic memory. Trends in Cognitive Sciences. 2022; 26(8):707–723. https://doi.org/10.1016/j.tics.2022.05.006.

Martin JT, Whittaker AH, Johnston SJ. Pupillometry and the vigilance decrement: Task-evoked but not baseline pupil measures reflect declining performance in visual vigilance tasks. European Journal of Neuroscience. 2022; 55(3):778–799. https://doi.org/10.1111/ejn.15585.

Massar SA, Lim J, Sasmita K, Chee MW. Rewards boost sustained attention through higher effort: A value-based decision making approach. Biological Psychology. 2016; 120:21–27. https://doi.org/10.1016/j.biopsycho.2016.07.019.

Mathôt S. Pupillometry: Psychology, physiology, and function. Journal of cognition. 2018a; 1(1):16. https://doi.org/10.5334/joc.18.

Mathôt S, Fabius J, Van Heusden E, Van der Stigchel S. Safe and sensible preprocessing and baseline correction of pupil-size data. Behavior research methods. 2018b; 50(1):94-106. https://doi.org/10.3758/s13428-017-1007-2.

Mathôt S, Aarts E, Verhage M, Veenvliet J, Dolan C, van der Sluis S, A simple way to reconstruct pupil size during eye blinks; 2013.

Mathôt S, Vilotijević A. Methods in cognitive pupillometry: Design, preprocessing, and statistical analysis. Behavior Research Methods. 2023; 55(6):3055–3077. https://doi.org/10.3758/s13428-022-01957-7.

McGinley M, David S, McCormick D. Cortical Membrane Potential Signature of Optimal States for Sensory Signal Detection. Neuron. 2015; 87(1):179–192. https://doi.org/10.1016/j.neuron.2015.05.038.

Megemont M, McBurney-Lin J, Yang H. Pupil diameter is not an accurate real-time readout of locus coeruleus activity. Elife. 2022; 11:e70510. https://doi.org/10.7554/eLife.70510.

Meissner SN, Bächinger M, Kikkert S, Imhof J, Missura S, Carro Dominguez M, Wenderoth N. Self-regulating arousal via pupil-based biofeedback. Nature Human Behaviour. 2024; 8(1):43–62. https://doi.org/10.1038/s41562-023-01729-z.

Merali Z. Error: why scientific programming does not compute. Nature. 2010; 467(7317):775–777. https://doi.org/10.1038/467775a.

Miller AL, Gross MP, Unsworth N. Individual differences in working memory capacity and long-term memory: The influence of intensity of attention to items at encoding as measured by pupil dilation. Journal of Memory and Language. 2019; 104:25–42. https://doi.org/10.1016/j.jml.2018.09.005.

Miller AL, Unsworth N. Variation in Attention at Encoding: Insights From Pupillometry and Eye Gaze Fixations. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2020; 46(12):2277–2294. https://doi.org/10.1037/xlm0000797.

Miller G, A scientist's nightmare: software problem leads to five retractions. American Association for the Advancement of Science; 2006. https://doi.org/10.1126/science.314.5807.1856.

Mittner M, pypillometry: A Python package for pupillometric analyses. Open Journals; 2020. https://doi.org/10.21105/joss.02348.

Murphy PR, O'connell RG, O'sullivan M, Robertson IH, Balsters JH. Pupil diameter covaries with BOLD activity in human locus coeruleus. Human brain mapping. 2014; 35(8):4140–4154. https://doi.org/10.1002/hbm.22466.

Murphy PR, Robertson IH, Balsters JH, O'connell RG. Pupillometry and P3 index the locus coeruleus–noradrenergic arousal function in humans. Psychophysiology. 2011; 48(11):1532–1543. https://doi.org/10.1111/j.1469-8986.2011.01226.x.

Oliva M, Anikin A. Pupil dilation reflects the time course of emotion recognition in human vocalizations. Scientific reports. 2018; 8(1):4871. https://doi.org/10.1038/s41598-018-23265-x.
Open Science Collaboration. Estimating the reproducibility of psychological science. Science. 2015; 349(6251):aac4716. https://doi.org/10.1126/science.aac4716.

Pedroni A, Bahreini A, Langer N. Automagic: Standardized preprocessing of big EEG data. NeuroImage. 2019; 200:460-473. https://doi.org/10.1016/j.neuroimage.2019.06.046.

Peysakhovich V, Vachon F, Dehais F. The impact of luminance on tonic and phasic pupillary responses to sustained cognitive load. International Journal of Psychophysiology. 2017; 112:40–45. https://doi.org/10.1016/j.ijpsycho.2016.12.003.

Piquado T, Isaacowitz D, Wingfield A. Pupillometry as a measure of cognitive effort in younger and older adults. Psychophysiology. 2010; 47(3):560–569. https://doi.org/10.1111/j.1469-8986.2009.00947.x.

Poe GR, Foote S, Eschenko O, Johansen JP, Bouret S, Aston-Jones G, Harley CW, Manahan-Vaughan D, Weinshenker D, Valentino R, et al. Locus coeruleus: a new look at the blue spot. Nature Reviews Neuroscience. 2020; 21(11):644–659. https://doi.org/10.1038/s41583-020-0360-9.

Poldrack RA, Baker CI, Durnez J, Gorgolewski KJ, Matthews PM, Munafʼo MR, Nichols TE, Poline JB, Vul E, Yarkoni T. Scanning the horizon: towards transparent and reproducible neuroimaging research. Nature reviews neuroscience. 2017; 18(2):115–126. https://doi.org/10.1038/nrn.2016.167.

Poldrack RA, Gorgolewski KJ, Varoquaux G. Computational and informatic advances for reproducible data analysis in neuroimaging. Annual Review of Biomedical Data Science. 2019; 2(1):119–138. https://doi.org/10.1146/annurev-biodatasci-072018-021237.

Provins C, MacNicol E, Seeley SH, Hagmann P, Esteban O. Quality control in functional MRI studies with MRIQC and fMRIPrep. Frontiers in Neuroimaging. 2023; 1:1073734. https://doi.org/10.3389/fnimg.2022.1073734.

R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria; 2025, https://www.R-project.org/.

Raasveldt M, Mühleisen H. Duckdb: an embeddable analytical database. InProceedings of the 2019 international conference on management of data. 2019 (pp. 1981-1984).

Reimer J, Froudarakis E, Cadwell C, Yatsenko D, Denfield G, Tolias A. Pupil Fluctuations Track Fast Switching of Cortical States during Quiet Wakefulness. Neuron. 2014; 84(2):355–362. https://doi.org/10.1016/j.neuron.2014.09.033.

Robison MK, Trost JM, Schor D, Gibson BS, Healey MK. Pupillary correlates of individual differences in long-term memory. Psychonomic Bulletin & Review. 2022; 29(4):1355–1366. https://doi.org/10.3758/s13423-022-02081-5.

Robison MK, Unsworth N. Pupillometry tracks fluctuations in working memory performance. Attention, Perception, & Psychophysics. 2019; 81(2):407–419. https://doi.org/10.3758/s13414-018-1618-4.

Rousseeuw PJ, Croux C. Alternatives to the median absolute deviation. Journal of the American Statistical association. 1993; 88(424):1273–1283.

Rousseeuw PJ, Hampel FR, Ronchetti EM, Stahel WA, Robust statistics: the approach based on influence functions. Wiley, New York; 1986.

Samuels ER, Szabadi E. Functional neuroanatomy of the noradrenergic locus coeruleus: its roles in the regulation of arousal and autonomic function part I: principles of functional organisation. Current neuropharmacology. 2008; 6(3):235–253. https://www.ingentaconnect.com/content/ben/cn/2008/00000006/00000003/art00005.

Schwartz ST, Nguyen KK, Yang H, deBettencourt MT, Tran TT, Madore KP, Wagner AD. Real-time reorienting of preparatory sustained attention lapsing during episodic retrieval using closed-loop pupillometry. In: 31st Annual Meeting of the Cognitive Neuroscience Society Toronto, Ontario, Canada; 2024. https://drive.google.com/file/d/1axi_MGuLCd6OG_XRAuM7_BratDAzmxcR/view?usp=sharing, conference presentation, https://drive.google.com/file/d/1rNRkWZMQCZfSoYS-JHWL2wB3Kh0UX4VZ/view.

Schwartz ST. eyeris: Flexible, Extensible, & Reproducible Processing of Pupil Data; 2025, https://doi.org/10.32614/CRAN.package.eyeris, R package version 1.1.0.

Schwartz ST, eyeris R package; 2025. Created in BioRender. https://BioRender.com/vkahl3u.

Schwartz ST, eyeris reports; 2025. Created in BioRender. https://BioRender.com/j2h2hzs.

Schwartz ST, Yang H, Xue AM, Wagner AD. Attending to Remember: Recent Advances in Methods and Theory. Current Directions in Psychological Science. 2025; https://doi.org/10.1177/09637214251339452.

Segal J. Scientists and software engineers: A tale of two cultures. In: Buckley J, Rooksby J, Bednarik R, editors. Proceedings of the 20th Annual Meeting of the Psychology of Programming Interest Group (PPIG 2008) Lancaster University, Lancaster, UK; 2008. https://oro.open.ac.uk/17671/4/PPIG_08Segal.pdf.

Simmons JP, Nelson LD, Simonsohn U. False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. Psychological science. 2011; 22(11):1359–1366. https://doi.org/10.1177/0956797611417632.

Smallwood J, Brown KS, Baird B, Mrazek MD, Franklin MS, Schooler JW. Insulation for Daydreams: A Role for Tonic Norepinephrine in the Facilitation of Internally Guided Thought. PLOS ONE. 2012 04; 7(4):1–5. https://doi.org/10.1371/journal.pone.0033706.

Šoški'c A, Styles SJ, Kappenman ES, Kovi'c V. Garden of forking paths in ERP research–Effects of varying preprocessing and analysis steps in an N400 experiment. Psychophysiology. 2024; 61(10):e14628. https://doi.org/10.1111/psyp.14628.

Spires-Jones TL, Hyman BT. The intersection of amyloid beta and tau at synapses in Alzheimer's disease. Neuron. 2014; 82(4):756–771. https://doi.org/10.1016/j.neuron.2014.05.004.

Steinhauer SR, Bradley MM, Siegle GJ, Roecklein KA, Dix A. Publication guidelines and recommendations for pupillary measurement in psychophysiological studies. Psychophysiology. 2022; 59(4):e14035. https://doi.org/10.1111/psyp.14035.

Steinhauer SR, Siegle GJ, Condray R, Pless M. Sympathetic and parasympathetic innervation of pupillary dilation during sustained processing. International journal of psychophysiology. 2004; 52(1):77–86. https://doi.org/10.1016/j.ijpsycho.2003.12.005.

Strauch C. The forgotten wave of early pupillometry research. Trends in Neurosciences. 2024; 47(8):571–572. https://doi.org/10.1016/j.tins.2024.06.002. Strauch C, Wang CA, Einhäuser W, Stigchel SVd, Naber M. Pupillometry as an integrated readout of distinct attentional networks. Trends in Neurosciences. 2022; 45(8):635–647. https://doi.org/10.1016/j.tins.2022.05.003.

Tsukahara JS, Harrison TL, Engle RW. The relationship between baseline pupil size and intelligence. Cognitive Psychology. 2016; 91:109–123. https://doi.org/10.1016/j.cogpsych.2016.10.001.

Tsukahara J, Pupillometry: An R package to preprocess pupil data (v0.8.1); 2020. https://github.com/dr-jt/pupillometry.

Unsworth N, Robison MK. Individual differences in the allocation of attention to items in working memory: Evidence from pupillometry. Psychonomic Bulletin & Review. 2015; 22(3):757–765. https://doi.org/10.3758/s13423-014-0747-6.

Unsworth N, Robison MK. Pupillary correlates of lapses of sustained attention. Cognitive, Affective, & Behavioral Neuroscience. 2016 Aug; 16(4):601–615. https://doi.org/10.3758/s13415-016-0417-4.

Unsworth N, Robison MK. The Importance of Arousal for Variation in Working Memory Capacity and Attention Control: A Latent Variable Pupillometry Study. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2017; 43(12):1962–1987. https://doi.org/10.1037/xlm0000421.

Unsworth N, Robison MK. A locus coeruleus-norepinephrine account of individual differences in working memory capacity and attention control. Psychonomic Bulletin & Review. 2017; 24(4):1282–1311. https://doi.org/10.3758/s13423-016-1220-5.

Unsworth N, Robison MK. Tracking arousal state and mind wandering with pupillometry. Cognitive, Affective, & Behavioral Neuroscience. 2018 Aug; 18(4):638–664. https://doi.org/10.3758/s13415-018-0594-4.
Urai AE, Braun A, Donner TH. Pupil-linked arousal is driven by decision uncertainty and alters serial choice bias. Nature communications. 2017; 8(1):14637. https://doi.org/10.1038/ncomms14637.

van den Brink RL, Murphy PR, Nieuwenhuis S. Pupil Diameter Tracks Lapses of Attention. PLOS ONE. 2016 10; 11(10):1–16. https://doi.org/10.1371/journal.pone.0165274.

Van Gerven PW, Paas F, Van Merriënboer JJ, Schmidt HG. Memory load and the cognitive pupillary response in aging. Psychophysiology. 2004; 41(2):167–174. https://doi.org/10.1111/j.1469-8986.2003.00148.x.

van Rij J, Hendriks P, Rijn Hv, Baayen RH, Wood SN. Analyzing the Time Course of Pupillometric Data. Trends in Hearing. 2019; 23:2331216519832483. https://doi.org/10.1177/2331216519832483.

Van Slooten JC, Jahfari S, Knapen T, Theeuwes J. How pupil responses track value-based decision-making during and after reinforcement learning. PLoS computational biology. 2018; 14(11):e1006632. https://doi.org/10.1371/journal.pcbi.1006632.

Varoquaux G. Beyond computational reproducibility, let us aim for reusability. IEEE CIS Newsletter on Cognitive and Developmental Systems. 2016; 13(2). https://inria.hal.science/hal-01592368v1, what is Computational Reproducibility?

Verney SP, Granholm E, Marshall SP. Pupillary responses on the visual backward masking task reflect general cognitive ability. International journal of psychophysiology. 2004; 52(1):23–36. https://doi.org/10.1016/j.ijpsycho.2003.12.003.

Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, et al. Welcome to the Tidyverse. Journal of open source software. 2019; 4(43):1686. https://doi.org/10.21105/joss.01686.

Widmann A, Schröger E, Wetzel N. Emotion lies in the eye of the listener: Emotional arousal to novel sounds is reflected in the sympathetic contribution to the pupil dilation response and the P3. Biological psychology. 2018; 133:10–17. https://doi.org/10.1016/j.biopsycho.2018.01.010.

Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten JW, Santos LBdS, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, et al. The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data. 2016; 3(1):160018. https://doi.org/10.1038/sdata.2016.18.

Wilson G, Aruliah DA, Brown CT, Chue Hong NP, Davis M, Guy RT, Haddock SH, Huff KD, Mitchell IM, Plumbley MD, et al. Best practices for scientific computing. PLoS biology. 2014; 12(1):e1001745. https://doi.org/10.1371/journal.pbio.1001745.

Wu J, Toporek A, Lin Q, Goldstein FC, Loring DW, Kelberman MA, Weinshenker D, Levey AI, Lah JJ, Qiu D. Probing locus coeruleus functional network in healthy aging and its association with Alzheimer's disease biomarkers using pupillometry. Alzheimer's Research & Therapy. 2025; 17(1):1–12. https://doi.org/10.1186/s13195-025-01701-1.

Zeileis A, Grothendieck G. zoo: S3 Infrastructure for Regular and Irregular Time Series. Journal of Statistical Software. 2005; 14(6):1–27. https://doi.org/10.18637/jss.v014.i06.

Zhao S, Bury G, Milne A, Chait M. Pupillometry as an Objective Measure of Sustained Attention in Young and Older Listeners. Trends in Hearing. 2019; 23:2331216519887815. https://doi.org/10.1177/2331216519887815.

Appendix: Custom *eyeris* Extension Source Code

1) Write core operation function logic.

```
winsorize_pupil <- function(x, prev_op, lower = .01, upper = .99) {
  vec <- x[[prev_op]]
  q <- quantile(vec, probs = c(lower, upper), na.rm = TRUE)
  vec[vec < q[1]] <- q[1]
  vec[vec > q[2]] <- q[2]
  vec
}
```

2) Create wrapper using the *pipeline_handler()* function.

```
#' Winsorize pupil values
#'
#' Applies winsorization to extreme pupil values within
#' each block.
#'
#' @param eyeris An `eyeris` object created by [load_asc()].
#' @param lower Lower quantile threshold. Default is 0.01.
#' @param upper Upper quantile threshold. Default is 0.99.
#'
#' @return Updated `eyeris` object with new winsorized
#' pupil column.
winsorize <- function(eyeris, lower = .01, upper = .99) {
  pipeline_handler(
          eyeris,
          winsorize_pupil,
          "winsorize",
          lower = lower,
          upper = upper
  )
}
```

3) Invoke custom function in *eyeris* pipeline call.

```
system.file("extdata", "memory.asc", package = "eyeris") |>
  eyeris::load_asc(block = "auto") |>
  eyeris::deblink(extend = 50) |>
  winsorize()
```