

## **Response to reviewers for**

**Manuscript Title:**  
**Analysis of human visual experience data**

**Submission Id:**  
**JOV-09808-2025**

11 December 2025

We thank the editor and the reviewers for their insights and feedback. We have made several changes to the manuscript based on this feedback. We reference these changes point by point, together with the reviewer's remarks. Reviewers comments are written verbatim, with our point by point replies in **bold**. We also incorporated other changes that were not raised by reviewers, but suggested by colleagues after reviewing the preprint. We have listed these points, and the respective adjustments, after Reviewer #2.

---

### *Reviewer 1:*

The manuscript describes the richness of the optical environment in terms of the human eye's exposure to light, particularly daylight. It then correctly identifies how software tools for data analysis are fragmented within the daylight/circadian research community, and proceeds to introduce the LightLogR package as a solution to this problem. That effort is duplicated in re-inventing analysis code, that errors go unseen, and that results from different studies are difficult to harmonise into larger datasets are factors which are too easily underestimated. As such, the existence of LightLogR is a treasure for the circadian community, and this manuscript is a welcomed advertisement for the package.

The remainder of the manuscript serves as a tutorial on how to import and analyse data from two example photosensors. The discussion mentions that the package is open source, actively maintained, and is open to feature requests and contributions from the community - these are all sound qualities that I like to see in software research tools. The authors are to be commended on what I'm sure was a formidable effort to bring together the code, documentation, tutorials and unit tests.

All that said, as a software developer myself, and as someone quite new to circadian research, I felt that the manuscript failed to sell me on why I should abandon my own software tools in favour of LightLogR. While ignoring any of the points below should not prevent publication, I think it would strengthen the manuscript if the following issues were to be addressed:

**Our reply:**

**We thank Reviewer 1 for the encouragement and constructive feedback. We have adjusted the manuscript according to the feedback and will break the changes down point by point**

**Reviewer 1:**

1) How difficult is it to add code for importing data from the light sensor used in my own research (assuming support does not already exist)? Is the output data structure of the import function well defined? Is there a template importer that I can copy and modify?

**Our reply:**

**There is a function in LightLogR that allows code injection to define an individual import function. That said, the import functions really don't restructure the raw data, at least if they are rectangular to begin with. That means that any rectangular dataframe can be loaded into R and used with LightLogR. What the import functions do is mainly about making sure there is a correctly formatted Datetime column, a grouping Id column, and an import summary. But it also offers a number of conveniences, like sorting by time, correcting for daylight savings time (if a device records over a period containing a DST jump, but does not adjust time stamps accordingly), setting cut-off dates, checking for duplicate rows, and more. We have extended the manuscript with the following paragraph and would also direct Reviewer 1 to the [supplement](#), which contains much more information on the import capabilities of the package.**

**New paragraph in chapter *Data processing summary***

**“Besides the `Clouclip` and `VEET`, LightLogR 0.10.0 contains import functions for 18 more wearable devices. The package further supports versions due to evolving data formats, and includes documentation for both [code-based](#) and [code-less](#) addition of new device import-functions.”**

**Reviewer 1:**

2) Having imported data from two (or more) heterogeneous sensor types (as shown in the manuscript), is there a common data format into which this data is imported, and does that format abstract away differences in sensor type? If so, I would be able to write a single dataflow for all the data from all of my heterogeneous sensors, and not maintain two (or more) parallel data flows which differ in trivially and irritatingly minor ways.

**Our reply:**

**In short, yes on the merge, no on adjustments of sensor types. As noted above, LightLogR's import does not change variable names of the raw data, nor does it calibrate device outputs. This means that a user would have to manually decide which variables of different devices they want to use and rename their columns to be identical. This can be done posthoc after importing**

from different devices, which is covered in the following paragraph we have added to the manuscript. The other way would be an import based on device type (and possibly other factors). An example of this can be seen [here](#). Once merged, a single pipeline can cover more than one device type.

### New chapter *Merging data streams*

„Note that while imports from different devices can be [merged](#), devices differ in their sensors, electronics, housing or diffuser form factors, and on-device data-processing pipelines. All of these factors affect the comparability of measurements, even when devices output the same variable (e.g., illuminance or distance). If data from different devices with the same measurement variable are to be merged, the corresponding variable names should be standardized beforehand - for example, renaming `Lux` and `LIGHT` to `illuminance`. If we wanted to analyse the `VEET` data together with the `Clouclip` data, for example, we would not have to rename anything, as both carry their illuminance measurements in the variable `Lux`. The following example shows how the combination of datasets would lead to a combined dataset, and how that would affect analysis outcomes. It is the responsibility of the researcher to perform device calibration and/or checks for a similar measurement fidelity.”

#### *Reviewer 1:*

3) How "rich" are the analysis tools available to imported data? Table 16 shows an example of subsetting the data to study "blue light", but what is the definition of "blue light" here - some physics-based definition based on the electromagnetic spectrum, or based on retinal photoreceptor sensitivity? If the latter, are there options for choosing different cone characteristics, like ala180 and ser180 variations of L-cone sensitivity?

#### **Our reply:**

The referred example is arbitrary in its definition for “short wavelength” or “blue” light, and the thresholds can be configured freely. LightLogR contains some action spectra (and weights) that are built into [spectral\\_integration\(\)](#), which are the alphaopic action spectra from CIE S026:2018 (containing cone, rod, and ipRGC sensitivities) and the 2° photometric standard observer from ISO/CIE 11664-1:2019. That said, the `spectral\_integration()` allows your own action spectra and weights, so implementing, e.g., different L-cone sensitivity functions would be easy.

We have added a [section](#) to the *Spectrum* analysis that showcases these defaults by calculating photopic and melanopic metrics from the spectrum.

#### *Reviewer 1:*

I could, of course, download LightLogR and explore the answers to these questions

myself, but it feels like a missed opportunity for the manuscript to promote these features to me. The addition of a few more explanatory sentences would be helpful in this regard. It may be sufficient, for example, to explain the dataCC structure when you first introduce it (line 264) - is this a universal structure? Could I rbind() dataCC with the structure imported from the VEET to produce a homogenous data set of spectra measurements, independent of the source instrument?

**Our reply:**

**We agree that these aspects should be better signposted, as the introduction of the data formats is done the supplement. We have made this now much clearer with an info box in the introduction on the scope and division of the article and supplement.**

*Reviewer 1:*

My other complaint with the manuscript was how the tutorial handled the behavioural data of the participants. I am shown code which handles periods of "visual breaks", "illuminance distribution", "average light exposure", etc, but no comment on whether the results of these analyses correspond to the sensor wearer's actual behaviour. Why, for example, does the participant in Table 9 have less mean daylight exposure at the weekend than during week days? For the majority of western-world professions, this is the opposite of what one would expect (and I'm left to assume that the participant was measured during a typical working week since there is no text to suggest otherwise). Some comment on what this participant does as a career (a landscape gardener who likes playing video games in her spare time, for example) would add useful meaning to the data in Table 9.

**Our reply:**

**We agree that the individual results deviate quite a bit from typical values seen in groups. We have added a paragraph to the scope to make readers aware of this:**

**"Lastly, the example data used in the article do not stem from a controlled experimental data collection but consist of pilot data gathered in an ecological setting without a fixed protocol. Given the substantial interindividual differences in visual experience metrics, and because the analyses focus on one participant at a time, the reported results should be interpreted as illustrative rather than representative of typical or population-level values."**

*Reviewer 1:*

Similarly, near-work episodes are of particular interest to the myopia community, but it's hard to assess the correctness of Table 7 when I don't know what the participant was doing during the measurement period (reading? playing outdoor sports?). And again, what was the participant in Table 5/Figure 1 doing such that they spent 60% of their weekend doing "Extremely near" work between 10 and 20 cm? Was this

participant extremely myopic? Or could this be a sensor calibration issue?

None of this is to question the correctness of the analyses in the tutorial - I'm merely asking for some sentences to provide helpful background context.

**Our reply:**

**We believe the previous answer covers some aspects of this point. However, we agree that this is a particular case in Table 5 / Figure 1. While the excess of very close viewing distances (~10cm) are in the data, i.e., the summary bins them correctly, it is confusing and might be an issue in the recording context. As prolonged viewing distances of 10 cm are not realistic, we have thus capped the analyses at 15 cm throughout that section.**

*Reviewer 1:*

Finally, I do not appear to have access to the GitHub site where the data from this tutorial is supposed to reside, so my apologies if the following is already implemented. Where the manuscript states "All necessary data and code are openly available

in the GitHub repository" is this the actual data used in the manuscript's tutorial? Could I use this to replicate the exact tables/graphs shown using the exact code shown? I have found such a sanity-checking process useful when exploring a new-to-me software package.

**Our reply:**

**That is correct – all the code and data to reproduce the article are found in this repository:**

[https://github.com/tsclab/ZaunerEtAl\\_JVis\\_2026](https://github.com/tsclab/ZaunerEtAl_JVis_2026)

**and hosted on GitHub pages:**

[https://tsclab.github.io/ZaunerEtAl\\_JVis\\_2026/](https://tsclab.github.io/ZaunerEtAl_JVis_2026/)

*Reviewer 1:*

Aside from those three complaints, I have a few minor suggestions and fixes:

I felt the authors had a slightly dismissive tone when describing vision experiments in "controlled laboratory settings" (lines 54-56). I appreciate the manuscript was trying to be concise, but I feel it is important to acknowledge that understanding phenomena in a controlled environment is an essential part of the scientific process, and, arguably, a necessary step before trying to interpret the chaotic real world.

**Our reply:**

**We agree that the tone was not appropriate, and it was not our intention to be dismissive of laboratory studies, which are our primary means of gaining a**

**mechanistic understanding. We have rephrased the paragraph so that it highlights the differences without attributing value:**

**"In controlled laboratory settings, light exposure can be held constant or manipulated parametrically. In contrast, real-world conditions are inherently complex and dynamic, and cannot be captured by single spot measurements. As people move in and between spaces (indoors and outdoors) and move their body, head, and eyes, exposure to the optical environment varies significantly ([Webler et al. 2019](#)) and is modulated by behavior ([Biller, Balakrishnan, and Spitschan 2024](#)). Wearable devices for measuring light exposure have thus emerged as vital tools to capture the ecological visual experience. These tools generate high-dimensional datasets that demand rigorous and flexible analysis strategies.**

**Reviewer 1:**

I assume "dataswert" on line 231 should have been "dataset"?

**Our reply:**

**That is correct, thank you for spotting the mistake**

**Reviewer 1:**

Table 14 caption (lines 693) would benefit from adding the word "illuminance" after "Longest period and total duration", to make a self-contained explanation for the table.

**Our reply:**

**We have rewritten the description accordingly**

---

**Reviewer 2:**

This article describes a tutorial to calculate a variety of visual experience metrics using LightLogR. Standardizing the description and analysis of visual experience data is important, as it will allow for more consistent comparisons across studies and improve the reliability of findings. Various definitions were presented in the tutorial, and it would be helpful if the authors could clarify whether these definitions can be modified, as no robust justification for their selection was provided. The presented analyses refer to data collected both within a single day and across multiple days. How have authors considered variability in the amount of data captured between individuals? Other comments are provided below.

**Our reply:**

**We thank Reviewer 2 for the encouragement and constructive feedback. We have adjusted the manuscript according to the feedback and will break the changes down point by point. Regarding your main point, we have added an**

info box to the *Introduction* on the scope of the article. This contains a short paragraph relating to the adjustability of definitions:

“To demonstrate the workflows, this article uses expert-informed definitions of metrics and metric parameters (see, e.g., Table 1, Table 2, and the non-wear detection rules based on activity data described in [Supplement 1](#)). These definitions and thresholds should not be interpreted as universal standards, nor are they hard-coded into the software package. For any application, parameter choices must be tailored to the research domain, study context and design, and the specifications of the wearable device. “

Regarding your second main point, on the variability of data captured between individuals. This aspect is actually covered in the [supplement](#), which covers all aspects relating to missing data. We have added another paragraph to the info box which reads:

“Further, the article is split up in the main analysis part, where all metrics are calculated, and the [Supplement 1](#), where data is imported, screened, and prepared. Thus, the reader is referred to [Supplement 1](#) for all aspects regarding data formats, preparation steps, and handling of gaps, i.e., missing data.”

In brief, only days with enough non-missing data are included in the analysis. This non-missingness threshold is also freely adjustable and depends on the analysis requirements.

Reviewer 2:

Line 71: please include references

Our reply:

We have added the two references to the line:

Hönekopp, A., and S. Weigelt. 2023. “Using Light Meters to Investigate the Light-Myopia Association - a Literature Review of Devices and Research Methods.” Journal Article. *Clin Ophthalmol* 17: 2737–60. <https://doi.org/10.2147/OPTH.S420631>.

Hartmeyer, S. L., and M. Andersen. 2023. “Towards a Framework for Light-Dosimetry Studies: Quantification Metrics.” Journal Article. *Lighting Research & Technology* 56 (4): 337–65. <https://doi.org/10.1177/14771535231170500>.

Reviewer 2:

Table 2: are there references to support distances used to define near and intermediate distance ranges? If total working distance is from 10 to 120 cm, what would 100-120cm be considered? Also, are there references to support definition for different time periods (continuous, interruptions, break etc)

Our reply:

The first and third question is covered with reference to the illustrative nature of the used metrics and definition, that we added with the info box:

**“To demonstrate the workflows, this article uses expert-informed definitions of metrics and metric parameters (see, e.g., Table 1, Table 2, and the non-wear detection rules based on activity data described in [Supplement 1](#). These definitions and thresholds should not be interpreted as universal standards, nor are they hard-coded into the software package. For any application, parameter choices must be tailored to the research domain, study context and design, and the specifications of the wearable device. “**

**There is however, mention in the text for some references to the use of these metrics in the *Methods and materials* section:**

**“In March 2025, two workshops with myopia researchers — initiated by the Research Data Alliance (RDA) Working Group on Optical Radiation Exposure and Visual Experience Data — focused on current needs and future opportunities in data analysis, including the development and standardization of metrics. Based on expert input from these workshops, the authors of this tutorial compiled a list of visual experience metrics, shown in [Table 1](#). These include many currently used metrics and definitions ([Wen et al. 2020, 2019; Bhandari and Ostrin 2020; Williams et al. 2019](#)), as well as new metrics enabled by spectrally-resolved measurements. While they are not derived by a formal consensus process, they are expert-informed and used in current scientific research, and thus will serve as example-definitions for metrics and thresholds throughout this article.“**

**Regarding your second question, regarding the total working distance of 10 to 120cm, we have added a footnote specifying that 120 is simply the upper measurement bound of the device.**

**Reviewer 2:**

Line 94: were data captured simultaneously from both devices or from two separate individuals/environmental conditions?

**Our reply:**

**We have added this information to the sentence, which now reads:**

**“We use example data from two devices (worn by different individuals and at different times) to showcase [...]”**

**Reviewer 2:**

Line 145: please confirm if distance and illumination data are captured simultaneously with the Clouclip

**Our reply:**

**We confirm this and have added the information to the specified line**

*Reviewer 2:*

Line 168: different wearable devices use varying notations to indicate sleep mode, have different measurement ranges and output files. Does this program have the capability to recognise and interpret data from all these different device types?

**Our reply:**

**The package contains custom import functions for currently 20 device types, with options to add more, either through a custom function or by providing a sample file to the developers. We have added a small paragraph to the section that reads:**

**“Besides the `Clouclip` and `VEET`, LightLogR 0.10.0 contains import functions for 18 more wearable devices. The package further supports versions due to evolving data formats, and includes documentation for both code-based and code-less additions of new device import-functions.”**

*Reviewer 2:*

Line 177: irregular or missing periods left as blank in export file?

**Our reply:**

**The way the analysis package determines missing or irregular data is based on a timeline of uninterrupted, equally spaced recording intervals from first until last measurement. If an observation does not fall on an expected timepoint in this timeline, or is missing completely, it is flagged as such. This is explained in more detail in the supplement, and the info box in the *Introduction* section refers to it more clearly.**

*Reviewer 2:*

Line 193: were there cases of irregular timestamps for the VEET dataset?

**Our reply:**

**Yes, there were and we have rephrased the sentence to reflect this more clearly:**

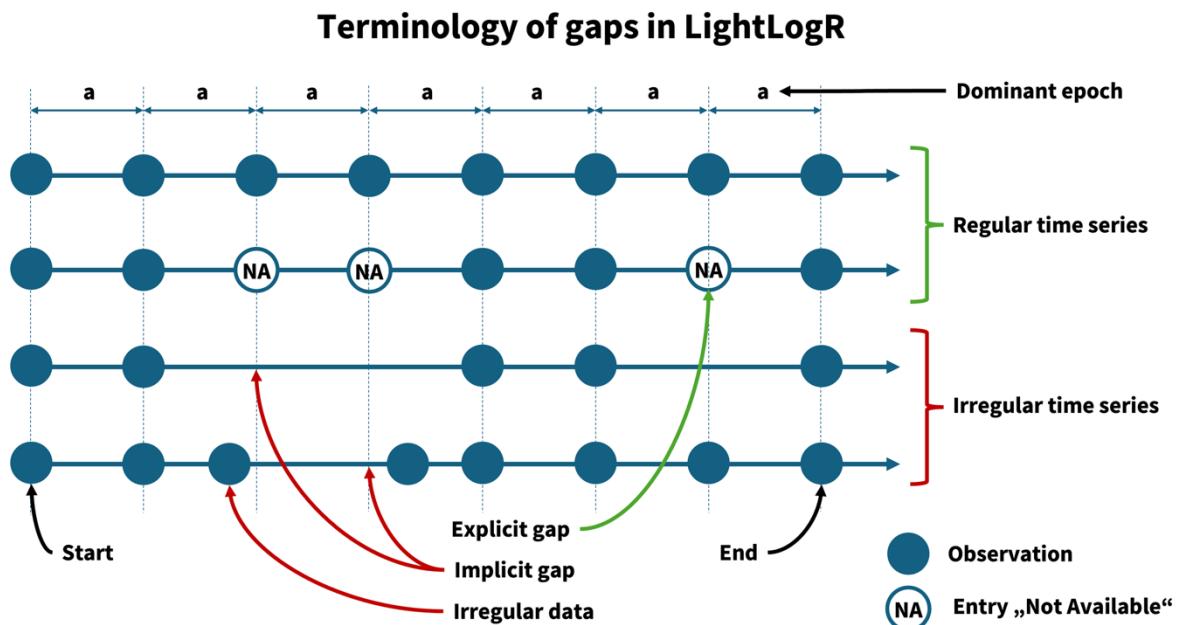
**“For the `VEET` device, data were provided as CSV logs (zipped on Github, due to size). We focused on the ambient light sensor modality first. Using `import\$VEET(..., modality = "ALS")`, we extracted the illuminance (`Lux`) data stream and its timestamps. The raw VEET data similarly contains irregular intervals and can contain missing periods (e.g., if the device stopped recording or was reset); the import summary flags these issues.”**

Reviewer 2:

Line 194: how did authors determine that missing data were actual implicit gaps?

Our reply:

Gaps in the package are based on an equally divided timeline of observations. An implicit gap is simply a missing observation at a timepoint when there should be one. The following figure explains the terminology that we have added to the supplement, where all preprocessing is detailed.



Reviewer 2:

Line 197: was there a criteria for minimum hours in a day to be considered as day data?

Our reply:

Not in this case. We have rewritten this line to address this topic, if only briefly:

"In our `Clouclip` example, days with <1 hour of recordings were dropped. This threshold should be adjusted based on how much complete days matter for a given analysis at hand. E.g., in circadian science, the metrics of interdaily stability and intradaily variation require measurements for each hour of the day."

Reviewer 2:

Line 205: how was data translated from 2 second to 5 second intervals? Is 5 seconds the proposed standard interval to use and if so, why?

Our reply:

We have added a sentence after the line reading

**"Aggregation was performed with the arithmetic mean of values in a 5-second bin."**

**5 seconds is not a standard, but it was the interval the Clouclip device used, thus we put them on the same interval. (this is stated in the previous line).**

*Reviewer 2:*

Line 206: how about data that had a late start or early finish, as that could also be missing hours of data at the beginning and end of day

**Our reply:**

**That is indeed correct. The gap-handling function fills those in as well if specified (which was the case). We have clarified this, the sentence now reads:**

**"We then inserted explicit missing entries for each whole day and removed days with more than one hour of missing illuminance data."**

*Reviewer 2:*

Line 2015: what is justification for 5 minute intervals?

**Our reply:**

**The justification we had is provided in the sentence:**

**"We aggregated the spectral data to 5-minute intervals to focus on broader trends and reduce data volume."**

**There is now justification beyond that – other intervals like 10 or 15 minutes would have worked equally well for our purpose.**

*Reviewer 2:*

Line 212: were the wavelength specific channels across standardized intervals between 415-940nm? What were the ranges of broad clear channels?

**Our reply:**

**We have expanded on the line, which now reads:**

**"[...] (approximately 415 nm through 940 nm, unequally spaced between 30 and 50 nm plus ,one broadband clear channel covering the whole range of individual channels and a dark channel) [...]"**

*Reviewer 2:*

Liner 236: consider using an alternative term to richer to avoid potential bias (here and throughout manuscript)

**Our reply:**

We have replaced the phrase throughout the manuscript, e.g., with “multimodal”.

*Reviewer 2:*

Line 322: justification/reference to support this definition of continuous near work?

**Our reply:**

**To stay in line with the illustrative nature of the provided metrics, as stated in the info-box, we have rephrased this sentence, which now reads:**

**“Continuous near-work can be understood as sustained viewing within a near distance for some minimum duration, allowing only brief interruptions.”**

*Reviewer 2:*

Line 360: what is the shorter minimum duration, and shouldn't there be a consistent duration for episodes, as it is then difficult to compare visual experience for example between weekend and weekdays, if the duration of the episode is different.

**Our reply:**

**We have rephrased this sentence to be less confusing, because while the minimum duration was changed for the example, it is consistent within the example and thus comparable. It now reads:**

**“This section extracts all near-work episodes (using a 5-second minimum duration to capture more routine near-work bouts) and summarizes three aspects:”**

*Reviewer 2:*

Line 388: does the same definition apply for consistent duration, where small breaks of < 1 minute are allowed?

**Our reply:**

**Yes, it would be possible to allow for these small breaks. The function arguments in the code cells show whether this is the case**

*Reviewer 2:*

Line 524: consider using another term than exemplary, to remove any bias

**Our reply:**

**We have replaced all instances of the term throughout the article**

*Reviewer 2:*

Line 527: is this simply because the person using the Clouclip did not spend much time outside? Is it possible to use data that included more variable lighting conditions, as it seems to then allude that the Clouclip is not capable of appropriating capturing lighting data.

**Our reply:**

**We agree that this could be misunderstood and have rephrased the paragraph:**

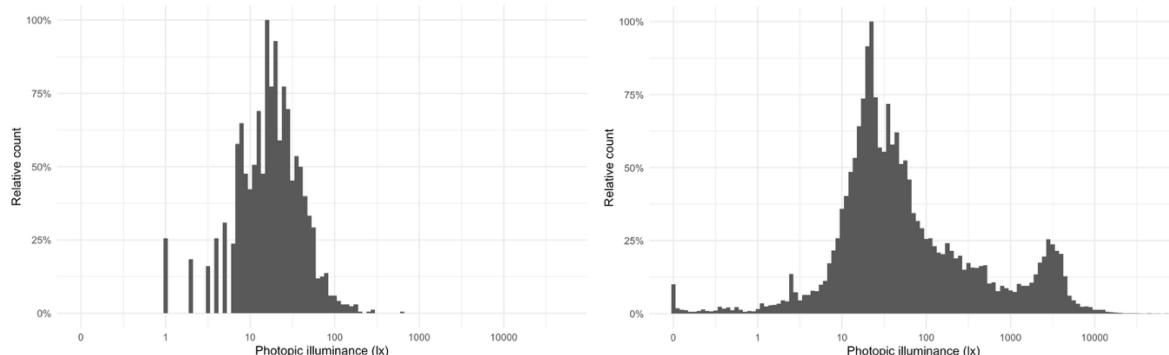
**“The ‘Clouclip’ illuminance data in our example cover indoor environments and are thus comparatively low, which would make certain daylight exposure summaries trivial or not meaningful. To better illustrate light exposure metrics, we turn to a different dataset, this one taken from the ‘VEET’ device’s illuminance data, which capture a broader range of lighting conditions (though both device types are able to capture broadly the same range of illuminance).”**

*Reviewer 2:*

Figure 6 and 7: please use same axes values to allow for comparison

**Our reply:**

**We have adjusted the figures.**



*Reviewer 2:*

Table 10: are the results presented here correct? Seems like the person was sitting in near darkness for most of the day if I am interpreting this correctly?

**Our reply:**

**It is merely that the high amount of zero values (during the night) brings the summary value down. In fact, the summary is more representative of the median than a classical average. That said, we have added a section detailing non-wear detection based on activity to the supplement and removed periods of non-wear from the VEET data. Accordingly, the distribution (Figure 7), and also the averages shown in Table 10 change.**

*Reviewer 2:*

Line 583: again, references/justification to support these definitions? Are these

correct, as I thought bright outdoor light was in the tens or hundreds of thousands of lux

**Our reply:**

**That is indeed correct. The light level at the eye, however, is much lower, as people avoid looking into very bright light sources if they can avoid it. As such, outside light exposure at the eye is much lower and usually in the region of a few thousand lux. We have added a paragraph to this section to elaborate on this aspect:**

**“While daylight levels can far exceed the recorded light levels, those are usually recorded with direct sunlight and without obstruction. Under normal viewing conditions, at eye level, and avoiding glare, daylight levels of a few thousand lux are at the higher end of the distribution ([Murukeshu, Zauner, and Spitschan 2025](#)). [Figure 7](#) shows a bimodal distribution, with the right mode representing outdoor lighting conditions. In a 2023 review of light dosimeters to investigate the light-myopia relationship ([Hönekopp and Weigelt 2023](#)), 1000 lx was the predominant cutoff value to distinguish indoor vs. outdoor environments. It is not, however, without critique, and both other thresholds ([Patterson Gentile et al. 2025](#)) and classification methods are proposed ([Tabandeh and Spitschan 2025](#)).”**

---

*Input from colleagues:*

[Minor] In 2.2, Table 2: Definition of Near work: Is 10-60 cm indeed the commonly accepted definition of near work in the Myopia research community? Is there a citation for this? It's sort of elaborated in the Distance ranges below, but both sides seem a little extreme:

- If I'm not mistaken, in order to focus at 10 cm you would already have something like -10 diopters of myopia, so I wonder if part of the range (like 10-15? 20? cm) should be re-labeled as likely sensor occlusion (hair, hand, etc.) rather than near work, and should be interpreted accordingly unless the research subject's actual refractive error indicates they can focus at such a short distance. I'm just concerned people might draw too much inference from the label...
- I've seen in other papers 50cm as the upper bound of near work. Is 60cm over-counting?

**Our reply:**

**With respect to the upper cutoff, it can indeed be debated whether 50 cm or 60 cm should be used to distinguish near from intermediate viewing distances. While 40 cm is commonly considered the clinical gold standard for near work**

and 60 cm for intermediate viewing in clinical settings, our study uses continuous objective distance measurements rather than discrete clinical categories. In this context, we and others have previously used 60 cm as the upper bound of near work in the objective myopia literature, and we have therefore retained this threshold. We agree, however, that the distinction is not absolute, and we have clarified illustrative nature of our examples in the revised manuscript.

Regarding the lower bound, a viewing distance of 10 cm does not imply a refractive error of -10 D, as participants are assumed to be corrected for distance; rather, it reflects an accommodative demand of approximately 10 D. While this level of accommodation may be theoretically possible in younger children, it is highly unlikely in adolescents or adults. Based on this consideration, and on our empirical observations, we now interpret distances below 15 cm with caution. In fact, we are currently treating only distances  $\geq$  15 cm as plausible working distances in our objective analyses. We acknowledge that we do not yet have a definitive explanation for the relatively high frequency of recorded distances below 15 cm, which may reflect sensor occlusion (e.g., hair or hand) or other measurement artifacts.

To reduce the risk of over-interpretation, we propose revising Table 2 (and corresponding text) to set the lower bound of near work at 15 cm rather than 10 cm, and to explicitly note that distances below this threshold are unlikely to represent feasible viewing distances. Overall, the chosen distance ranges reflect a combination of theoretical considerations, empirical observations, and practical constraints inherent to objective measurement of viewing distance, and we have revised the manuscript to make this clearer to readers.

*Input from colleagues:*

**In 3.1.1 Definition of 3.1.1 Total wear time daily:** I'm concerned about the accuracy of the definition "amount of time the device was actively collecting distance data each day (i.e. the time the device was worn and operational)" if the way it is derived is "summing all intervals where a valid distance measurement is present, ignoring periods where data are missing or the device was off".

Using VEET as an example, a distance measurement would be present every 2 seconds whenever the device is operating, even when the device is plugged into the wall (charging) and obviously not being worn, or when it is momentarily not being worn but pointed at a near object. As long as the battery was not allowed to run down then the "total wear duration" by this definition would be up to 24 hours each day. Conversely the "ignoring periods where data are missing" by looking simply at the detection of objects from the TOF alone might under-count wearing time — i.e. if the user is staring at far-away objects beyond the TOF sensor's range (e.g. watching a sunset or a baseball game).

I think for this to be a useful measurement that could be used for further analysis, an attempt should be made to remove obvious non-wear time.

If that's out of the scope, then perhaps it should be renamed to "Total recording time" so as to not over-promise what it does, until there is a metric that has higher fidelity to actual user wear time.

**Our reply:**

**This is an important aspect. We have added a section to the supplement that explains how non-wear detection works and implemented an example based on the IMU, which is then used to remove non-wear times from the example ALS and TOF modalities. With this reduced dataset (set non-wear to missing), the analysis is more robust**