Peer Review File

Article information: https://dx.doi.org/10.21037/atm-20-6440

To the editor: I'm grateful for these thoughtful comments and have made every effort to address each one. Throughout this document, normal and bold-faced fonts are used to reproduce the original reviewer comments, while my responses are in italics.

Reviewer A:

General comments:

This manuscript reviews key developments towards an 'optoretinogram', or noninvasive optical measurement of retinal function, with specific emphasis on the past decade or so. Overall, it is a fairly well-written and thorough review, but it could be improved substantially with some alterations. The authors enthusiasm and grasp of the key developments for the more recent work is evident throughout those sections and those read very well. However, portions of the manuscript that review the earlier experiments could be altered to improve the readability (see several specific comments below). Numerous findings are summarized in that section, along with specific measurements, magnitudes and polarities, as are hypotheses about underlying mechanisms and it would be useful to add a table or some other element into that section to summarize/compare the key early findings. It would also help to keep track between the different methods and animal species samples were derived from, etc. And really it would help to highlight the key findings, as at points these sections read like lists of facts without linking prose to tell a story about why these important figures or values are being highlighted. The figure provided is very helpful for summarizing and organizing the measurements made with the modern systems but it would be nice to see something for the earlier work, a timeline perhaps. In terms of overall readability, there is also some lack of transitional material between the different sections, so the flow and transition between each section should be reassessed carefully (some specific comments below). Finally, I think that there should be added an overall conclusion or summary at the end of the manuscript. At present, the report seems to end rather abruptly, and I was left looking for the conclusion. It would be nice to have some forward-looking statements at the end, perhaps as a bulleted list or short paragraph. Additional specific comments towards improvement are provided below.

Specific comments:

Lines 19-28: The author should consider consolidating these two short paragraphs into a single paragraph. Perhaps consider rewriting the first paragraph to focus on the key points made in these two paragraphs: 1) human vision begins at photoreceptors, 2) vision loss in retinal disease is most commonly due to loss of retinal neurons 3) present methods for testing retinal neural function and their strengths/limits

Agree

Line 20: The phrasing "which is then filtered and compressed" is awkward and doesn't fully capture the complexity of postcentral neural computation. Please consider rewording.

Agree: has been changed to the more general "post-processed"

Line 26: The author states "... methods have been used to learn *much* about..." It would be nice to provide a few examples or maybe drop this sentence.

Agree: citations have been aded to reviews of clinical research done using psychophysics (viz., perimetry) and several types of ERG.

Line 29: The author seems to mix both clinical and experimental testing methods together here when speaking about the psychophysical and electrophysiological methods. If it was limited to only discussion of clinical methodologies, I would fully agree with the statement that these are "... fundamentally blunt methods...". However, the recent work towards single-cell psychophysics (e.g. Roorda and collaborators)

would beg to differ with that assessment. I think the author should either stick to clinical measurements and their strengths/weaknesses here or acknowledge and cite some of the single-cell psychophysics work.

Agree: added the qualifier 'clinical', as well as citations to single-cell psychophysics.

Lines 56 & 57: Since this is not an ophthalmology specific journal, the authors might consider adding a paragraph or a figure somewhere around here summarizing the key retinal anatomical structures that will be discussed.

Agree: a figure explaining the basic anatomy and neural pathway has been added.

Line 102: It is not clear what the significance of the different temperatures listed here is with respect to the frog. I imagine frogs must be capable of vision at lower temperature levels than human body temperature so the relevance of 35C for the frog is not entirely clear. Here and elsewhere it would be best to only include the relevant details, or to omit the value of the measurement if the actual value is not important for the point you are trying to make. For example, if the point is that it was not observed at human body temperatures but only at lower temperatures then you could state something along those lines.

Agree. The sentence was removed.

Lines 103-104: Here and elsewhere in the manuscript the tense changes. The author should ensure that this is consistent throughout the manuscript to improve readability. Here within the paragraph "They also reported..." changes to in the next sentence "They attribute the...".

Agree. Past tense has been used throughout.

Lines 104-106: The author should consider rephrasing several of the statements that contain the phrasing "..., which..." to "that" to improve readability. For example, that convention is used here three times in a row where using *that* would be cleaner: "...diffraction peaks, which suggested disorganization..." can change to "...diffraction peaks *that* suggested..."

Agree. This has been rewritten.

Intro, lines 109-115: This paragraph should be rewritten. The first sentence is very long its second clause does not seem logically consistent with the second sentence. Also, please consider avoiding starting a sentence with "Of course,...".

Agree. This has been rewritten.

Intro, lines 128 and elsewhere: There are enough acronyms here already, please just write out positive instead of denoting these as p and n.

I agree that there are many acronyms already, but I hesitate to replace the abbreviations from the literature with my own words, because these signals–the "p" and "n" signals–were not described as "positive" and "negative" in the original papers. I have added "positive" and "negative", but left the original descriptions in place as well.

Intro, lines 129-131: This is an important point worth highlighting, you might want to state explicitly here the link between the scattering change and the chemical change. These paragraphs in this section can at times feel like enumerated lists of facts and it would be useful to have paragraphs that have some summary sentences at the end of them to highlight the key information.

Agree. Wording was changed to reflect this connection.

Intro, lines 132-137: Please revise this section to avoid having these two disjointed one sentence paragraphs here.

Agree. Done.

Intro, lines 144-145: This convention of starting a sentence with the authors names is somewhat overused throughout. Perhaps limit that convention to just the most key findings?

Agree. I have attempted to reduce these.

Intro, line 161: Consider starting a new paragraph after "... readjustment of the ROS."

Agree. This too-long paragraph was split into three.

Intro, lines 164-167: This sentence is awkwardly worded, please consider restating.

Agree. Reworded.

Intro, lines 174-175: The last sentence of this section is awkwardly written and does not serve as a good summary or transition between sections here. Please consider rewriting.

RSJ: REVISIT THIS

Intro, lines 178-180: The sentence that begins "Where these efforst were directly.." is difficult to follow, please rewrite for clarity.

Agree. Reworded.

Intro, lines 226-227: The author states here "... likely due to the cones." It would be useful to provide just a little more information here on why it was found to be likely due to the cones.

The reasons for it being likely the cones–dark adaptation duration compared to the magnitude of the reflectance changes–is not of great interest. I'm changing it to 'photoreceptors' instead.

Intro, line 231 and in many other locations: The author uses the word 'viz.' here and elsewhere. This is a very uncommon word and I found it off-putting. I suppose it is more common in British English but as an American reader it is not a word I encounter often, especially not in scientific writing. I have only ever usually seen it used in legal/judicial writing in the US. My recommendation would be to remove it throughout to improve readability.

Agree. Replaced with 'namely' throughout.

Intro, line 235: The author is referring here to his past work here. Elsewhere he refers to his work in the first person but here it is in the third person. I think all references to the authors own work should be written in first person to be consistent, with the specific paper referenced at the appropriate spot.

I did not find a place where I used a first-person pronound in the context of my own work; I believe I used "they" throughout. Nevertheless, I've replaced this with "we".

Lines 238-240: The reason for this response dependence on optical path length may not be readily apparent to all readers. It would be helpful for the author to describe why this is the case succinctly here.

Agree. I added two sentences describing the principle of coherence and the meaning of coherence length. I do not believe it is overly technical, and that it clarifies this result.

Lines 256-258: This is awkwardly phrased, consider rewording. A visual aid might be useful here as well so that the relevant parameters could be compared graphically.

Agree. A figure explaining the relationship between coherence length and interference has been added.

Line 284: insert the word 'segment' between inner and outer here.

I agree in principle that 'inner segment-outer segment' is more in keeping with the literature, 'innder-outer segment' is sufficiently clear, and slightly more concise.

Lines 296-298: This sentence is phrased awkwardly, please rephrase for clarity.

Agree. Rephrased

Line 298-300: This is an important point that is somewhat overlooked as you jump between frogs, cows, primates, etc. I would recommend reiterating this point elsewhere when necessary.

Agree.

Line 326: The authors phrasing here implies that a clinical SD-OCT can distinctly delineate the COST. This does not appear to be the intended convention. When referring to clinical measuremnets where the ROST and COST might not be distinctly visible, it would probably make sense to stick to clinical terminology.

Agree. This has been clarified in the manuscript.

Line 342: This is written strangely, shouldn't there just one approximate sign do for the whole range? (i.e. ~50–400nm)

Agree. Fixed.

Lines 346-350: it is not clear how this short paragraph logically links to this section (aside from using OCT). It would be useful to have a summary at the end of each of these sections.

Agree. To address this, the section has been divided into two subsections, one on scattering changes and the other on length changes. Hopefully this unifies the respective experiments.

Lines 347-348: Please provide the details on the Spectralis device or just omit the brand name if it is not relevant.

Agree. Done.

Lines 363, 372, 377 and 378: The same phrasing "_____ recognized,. . . ."is repeated four times here in short succession (and some more times later in this section). Consider rephrasing to improve readability.

Agree. Done.

Lines 400-401: The phrasing used here is awkward, please consider stating 'not in foveal or parafoveal cones' before 'likely owing to. . . .". Also, is there a relevant distinction to be made here about foveal vs. parafoveal cones?

Agree. The text has been changed.

Lines 411-412: Rather than stating 'nominally in the fovea' perhaps you could state 'just outside the fovea', since most anatomical descriptions set the fovea to be ~1 mm (~3 degrees) in diameter. Alternatively, you could use the term parafovea.

Agree. Fixed

Line 428: I think this overstates what has been done at present. The 'complete' rod mosaic is rarely imaged and has only been accomplished in a small number of eyes at some specific eccentricities.

Agree. The text was modified to make this clear.

Line 439: Why is it important to cite the number of photoisomerizations here? Please add some relevant detail.

Agree. Text changed to describe fraction of photopigment bleaching.

Line 440: Why would this be premature? Please elaborate.

I believe it's premature because we don't know whether the ORG signals should have kinetics similar to electrophysiological responses of rods and cones. Text was added to clarify.

Line 480: Were the isomerizations here estimated? Please elaborate.

The reviewer is correct. The language was changed to photopigment bleaching fraction.

Line 484-487: It is clear that the author is trying to assign attribution here to Sabesan for the term optoretinogram, however, the earlier work appears to have 'coined the term'. So, it is not clear to me that it makes sense to specifically call out Ram Sabesan by name here. I would leave this up to the editor to determine how to handle this.

I (reluctantly) agree. It feels unfair to take credit in publications (Azimipour, 2020) for the term, when I definitely borrowed it from Ram's conference presentations. I have tried to make this clear without mentioning Ram by name.

Lines 541-635: this section really needs to address the fact that it is possible that these measurements could be difficult to make in certain disease states. It would be useful to comment on some pathologies and how these might affect the measurements (e.g. drusen and pigmentary changes in AMD, lipofuscin precursor accumulation in cones in Stargardt's disease, etc.) There are numerous reasons why measurements that can be performed on laboratory observers in a bite bar (or some kind of head restraint system, such as has been shown for SS-FF-OCT) might be impossible to do in patients. The author should highlight some of the limitations and challenges to overcome to allow this work to move to the clinic. This is particularly important given the title of the review...

Line 547: This is really speculative here and I would posit that it would depend on the particular disease and also on the stage of disease. These measurements have not been validated in disease states and in some diseases they may be difficult or impossible to measure due to retinal pathology.

I agree with both of the comments above. A new section "Realizing the ORG's clinical potential" has been added, with acknowledgements of the challenges posed by the needs of patients (and clinicians).

Lines 581-582: I think that this needs to be toned down. You should state something along the lines of "...thus should be regarded skeptically until replicated." Or something like that.

I appreciate this comment and have made this change.

Lines 601-602: A sentence is needed here to summarize the preceding paragraph and put the findings in context.

A sentence has been added to provide context for this paragraph

Lines 602-611: It would be useful to expand on this section and provide some justifications for the proposed predictions for clinical applications here.

I appreciate the comment, but my expertise on the clinical applicability of the ERP is limited. I have reordered this section to emphasize the elongation phase and the importance of phototransduction changes for clinical translation.

Lines 62-624: This paragraph seems out of place here. Consider reordering.

A sentence has been added to provide context for this paragraph

Lines 686-687: Again here, it would be useful to have a paragraph that summarizes the essential points laid out in the preceding paragraphs of this section.

The preceding paragraphs have been rearranged in such a way to provide more context.

Line 700: It might be important to point out here that Grieve and Roorda pooled the measurements across hundreds of cones.

Added when the Grieve paper was introduced.

Line 708: I think that the author should not abbreviate optical path length here as OPL. There are already too many abbreviations for the non-expert reader to keep track of and this one is a problem even for those with vision relevant backgrounds as many ophthalmologists with automatically associate "OPL" with outer plexiform layer.

Agreed!

Line 746: The author should explain what "chirped" means for those readers that may not be familiar with that term.

Agree. Changed.

Lines 759-760: There is a need to add a summarizing paragraph here as well – these could summarize the preceding and also strengths/weaknesses of each measurement approach.

Agree, the text was modififed to describe the strengths and weaknesses of this approach.

Line 770: The author should note here or somewhere in this section that these computational and storage limits will likely be obviated with advancements in computing technology.

Agree. This sentence was at the bottom of the section, but I moved it upward to anticipate this question.

Line 771-774: Please consider rewriting this very long sentence that is also a paragraph to improve readability.

Agree. Changed.

Reviewer B:

I greatly enjoyed reading this very comprehensive review of the up and coming field of optoretinography. I have some suggestions and comments to improve the paper.

1. Angular dependence of the scattering response is mentioned in a few places (lines 118, 152, 628) and is quite interesting. Could waveguiding by photoreceptors increase the on-axis response, and would it decrease if photoreceptors were misaligned due to pathology? Is the on-axis response consistent across retinal eccentricities and could it be affected by the Stiles-Crawford effect? Is there an angular scattering difference between rods and cones? These seem like interesting questions to address in future even if they do not yet appear in the literature.

   This is a very interesting comment, but I feel that it is outside of the scope of this review to discuss disease-related changes in opthlamic images other than functional images.

2. The paper talks of responses from retinal neurons but is almost entirely focused on photoreceptors. At one point a ganglion cell signal is mentioned from the FFSSOCT approach. Why does only FFSSOCT detect RGC responses – because other techniques are focused only on PR layers or because they have searched but do not detect them? What could be expected in patients with degenerated or absent photoreceptors when using this technique? Could ORG be useful in patients undergoing regenerative therapy, for example optogenetics where other cell layers may be light sensitive? Maybe a comment could be made on some of these issues somewhere in the paper.

   The reviewer makes a great point. A paragraph has been added explaining the early interest in photoreceptors.

3. The sections detailing the various in vitro and in vivo results from different teams are very nicely explained in the text. However it is a little hard to keep track of what percentage change, or length change, or time course, was seen with each technique as they are so numerous. The figure nicely summarizes the principle sources of changes seen but I feel this could be complemented by a table summarizing the findings of different groups, with columns headed eg reference and date, method, % change, location of change, length change, time course, species, in vivo/in vitro, etc.

   I agree with the reviewer, but after spending weeks trying to reconcile various measurements I finally gave up. I suspect many of the differences are due to unpublished methodological differences, and thus would be difficult to reconcile in any meaningful way.

4. Lines 346-51: a study is mentioned which sought to observe effects in Alzheimer's patients. Only the results in control subjects are discussed here. Was anything observed in the patients?

   A sentence summarizing the Alzheimer's patients' response was added.

5. When presenting FFSSOCT the authors state that the method has limitations but does not detail what these are – it could be helpful just to add a short phrase here to explain.

   I have added a sentence describing these limitations. I hope it is not too technical for this journal.

6. I think there is a word missing or a word too many in the phrase at line 499 – maybe "that have translational potential" should be just "have" without the "that"?

Nice catch! Fixed.

7. From line 501 and figure legend – there is an odd formatting error in the pdf I received where the font changes and then is followed by what I understand to be the figure legend but without a figure numbering, then on the next page we return to regular font. Just a little confused by the formatting on where the figure and legend fit in relation to the list on page 501.

   I don't know how to fix this, as I think it arose in generation of the PDF.

8. In the figure, the cone and rod are shown at very similar size – I guess this is realistic at some eccentricities but it may be more visually obvious which is which if the cone were drawn a little fatter and the rod a little longer.

   I'm reluctant to change the figure because the revision is already much delayed. To the editor: if you think this is important, please let me know and I'll change it.

9. Line 606 – why 25%?

   The reviewer is correct. This is unnecessary, and removed.

10. Some confusion with acronyms rod OS, ROS, OS, etc, especially at line 713. Check for consistency throughout the paper and for first definitions.

    Agree. This has been fixed.

11. Line 765: data volumes are large rather than high

    Agree. Fixed.

12. A general comment is that the paper is entitled "Toward a clinical ORG" but as basically all work so far has been in healthy normal eyes, I understand that the clinical data is of course very sparse to non-existent. Two mentions of studies seeking to look at ORG in pathology are made in lines 780-785 but are hidden within a longer paragraph that includes information on animal studies. It would be nice to add a stand-alone conclusion paragraph mentioning the earliest clinical data that is available and discussing when we may see clinical use of ORG and what we might expect.

    Agree. Several paragraphs have been added detailing the challenges of clinical translation.

Reviewer C:

This review describes light-evoked optical changes in neurons including photoreceptors, extending from the classical measurements of light scattering to the more recent phase-resolved measurements in OCT. It sheds light on putative mechanisms of the optical changes, their relevance for clinical and basic science exploration and the constraints on hardware/software for various applications. Overall, the review is very comprehensive. It does an especially remarkable job in distilling decades of experiments from ex vivo, and in vivo animals and humans work and drawing comparisons between all. My comments and suggestions are aimed mostly at clarity, mechanistic underpinnings, completeness and striking a balance in the interpretations presented in the review.

First and foremost – The purpose of a review in my opinion is to summarize the current state of the understanding of a field. It is true that the area of optical imaging of neural activity in the living retina, aka optoretinography (ORG) is rapidly developing, but mainly in the area of hardware and technology. However, the understanding of their underlying mechanisms is currently very lacking. Explicitly stating that the technology developments are far leading the study of the ORG mechanisms is extremely critical for the reader to assess the current state of the field and this would make most sense in the abstract or early on in the review. If this was stated already, I apologize for overlooking or it just needs to be more explicit.

This is a great point, and I have added it–using the reviewer's nearly exact wording–near the beginning of the manuscript.

Line 12 and elsewhere: On nomenclature, the review uses the terms "elongation", "deformation" and others that refer to a "physical" change in the neurons, in this case the outer segment. Without a tool such as X-ray

diffraction, or serial electron microscopy, one cannot confirm the optical phase changes that have recently been observed are indeed attributable in their entirety to a "physical" change, and that either changes in refractive index or outer segment length may be responsible. Without a loss in the message, one could consider "activity" or something vaguer since this remains unknown. An alternative is to lay out your specific definition of "elongation" or "deformation" in the beginning and use it throughout.

A sentence has been added describing the use of the term "deformation", as an umbrella for contraction and elongation.

Introduction- Lines 19 – 42 has just one citation. The reader would appreciate more references to work on ERG and psychophysics. Related – even though the review does not deal with blood flow, readers may want to see just a mention of blood flow in the intro as a "mainstay" of functional neural testing – Line 23

A sentence has been added distinguishing optoretinography from "functional imaging" a la Grinvald's functional imager.

Perimetry – Line 39, For Goldmann perimetry, the stimulus size ranges from 6 arc-min to 2 deg, i.e. the precision can be substantially smaller than 1 deg., perhaps limited ultimately by eye motion and the eye's optics. Now, commercial fundus-guided perimeters are common, that would improve the limit further. Custom-made adaptive optics (AO) microperimetry with eye-tracking are now present in some labs that allow psychophysical testing at the scale on single and multiple cells. Subjects' input and localization of visual pathways still remain challenging, though fundus-guided microperimetry and AO microperimetry would be valuable additions for completeness on the state-of-the-art for current methodologies of functional visual testing in the clinic. Some very relevant papers on the topic: PMIDs : 22446720, 25587056, 33022378

The reviewer is correct about Goldmann spot sizes and microperimetry spot sizes. However, the smallest Goldmann spots are rarely used, and the standard clinical reports from fundus-guided microperimeters like the Maia still offer poor spatial resolution, possibly due to the constraints of eye motion combined with the relatively long (200 ms) stimulus duration. I have changed the language to be more generic, and added citations to AO microperimetry.

Line 136: Not clear how a change in birefringence is consistent with a change in OS disc spacing.

I hesitate to go any deeper into this issue, but have added a second citation (Aguirre) which describes the effects of disc spacing on phase retardation. I also clarified that this is transverse birefringence, in case the reviewer thought I was referring to on-axis imaging.

Line 147: Is it known that bovine photoreceptors express a ratio of 8:1 between opsins and transducin ? There are known to be differences between species and in the types of photoreceptors (rods vs. cones). Please check. If such data is not available, I'd suggest stating the caveat.

One paper says 8:1 and another says 10:1–the 10% difference is easily attributable to untested assumptions in each, but I don't want to belabor the point, since both roughly confirm the idea that R transducin interactions limit the scattering changes. I've cited Kahlert and Hofmann.*

Line 183: typo? Hunter, Merigan and Schallek ? Missing Merigan

Corrected.

Since in Line 194 and 195, the definition of scattering and volume changes are stated, some care must be taken to be consistent with these definitions. For instance, it is not clear which of the two in Line 200, light-evoked "changes" refer to. Also, the reader might expect densitometry to be discussed further in the review since some classical work on light-evoked scattering, and changes in general, in photoreceptors has its foundations there (though my own view is consistent with that of the authors to exclude densitometry). One way to set the reader's expectation would be to state that you refer mainly to near-infra-red wavelengths (except perhaps Bedggood & Metha), where absorption is extremely low (But see PMID: 25316726)

Agree. I have changed the text to better guide reader expectations.

Lines 235-258, the coherence lengths of all other studies except the range used in Jonnal et al. 2007 are indicated here. Could this missing coherence length be included as well for completeness and comparison

with the other studies?

Agree. A figure has been added describing the issue of coherence length in more depth.

Line 243 : "observed elongation" – No elongation was observed. Oscillations in intensity were "observed" which led to a hypothesis that the oscillations could result from interference and a change in optical path length. Please edit for accuracy as Line 392 says elongations measured for the first time.

Agree. This has been changed.

Line 265, "onset of the OPL changes occurred within 2ms" : Indicate that this is 2ms after stimulus onset. Same sentence : not sure how a fast change indicates an association of phototransduction. Do you mean, in comparison to the slower intensity changes observed in disc shedding and renewal ? Suggest explicitly stating for the naïve reader.

Agree. The text has been changed.

Line 332 – did the responses "peak" but also "saturate" at 2 min after bleach ?

I'm not sure if I understand this question. As I understand it, responses peak as a function of time and saturate as a function of irradiance.

Line 337 – The reader would find the axial resolution and other post-processing steps of this study useful since they were measuring a few 100s of nm changes in the retina, with amplitude alone and not phase.

Agreed. A sentence describing the axial resolution and role of averaging has been added.

Lines 363-365 – some more insight into why phase differences would not suffer from eye movement artifacts would be valuable. In addition to eye movements that create larger phase shifts than the movements of interest, there are other factors – air currents, vibrations and so on that affect small phase measurements. It would be helpful to add a short discussion on how common mode noise associated with these factors are effectively removed with referenced phase measurements.

Great point. Added.

Line 382-389 – it is not clear what the upper limit for the magnitude of higher order aberrations correctable with digital aberration correction is in full-field OCT ? This is important to compare against the later adaptive optics studies.

I agree, but these limits have not been fully explored. I have added a couple of sentences describing the computational tractability and limitations by the OCT SNR.

Line 396 and 413 and 420 : The observation – dose-dependent variation in elongation velocity – is mistakenly attributed to Azimipour et al., Optica 2019. This below is the full description of the pertinent results in the paper which has no mention of the rate of elongation in cones: "It is apparent that the magnitude of elongation varies with stimulus strength, as does the time required for the cone to recover its baseline length. This observed elongation of foveal/parafoveal cones, and the dependence of elongation on stimulus intensity, is qualitatively consistent with previously reported elongation of peripheral cones using full-field OCT with computational aberration correction"

This may have been qualitatively observed in this paper, but no mention of it in the results of the original paper and no quantitative description make it impossible to attribute this observation to it. It is strongly insisted that this be corrected and it be duly noted that the dose-dependent elongation "velocity/rate/slope etc" was first observed and quantified by Zhang et al. PNAS 2019. This Zhang et al. 2019 paper used a novel form of analysis where they reported the phase at the level of B-scans. This gave them high temporal resolution of 3 kHz (~ 300 microseconds) at the cost of spatial resolution to quantify the elongation velocity in their paper. In contrast, Azimipour et al. 2019 had 100 times lower temporal resolution of 30 Hz, inadequate to quantify the rate of elongation vs. stimulus intensity in cones. This speed is sufficient for rods as it was later shown in the Azimipour et al. 2020 paper.

I agree that the 2019 paper does not state this. However, the data shown in the paper show clear evidence. I have changed the wording to avoid misinterpretation.

Lines 401-404 – The reader would benefit greatly from an expanded discussion of RGC layer in addition to photoreceptors. In general, the focus in the review is disproportionately higher on the latter.

I agree with the reviewer's sentiment, but little work has been done on the topic. Moreover, to give RGCs the same historical treatment as photoreceptors would double the length of this review. I feel that it is out of the scope of this paper.

Line 406 – This is incorrect. An A-scan rate of 1.6 MHz is not the fastest AO-OCT system to-date and still faster AO-OCT systems have been published – Eg : Lines 476 and 494. When the paper was published, perhaps it was the fastest, so one could say "fastest AO-OCT at the time" or similar.

Agreed. The text was changed.

Line 408 – 2.5 deg is a very large estimate for the size of the fovea. Some might argue the rod-free zone is what constitutes the fovea and it is much smaller than 2.5 deg. Insist changing "foveal photoreceptors" to "parafoveal".

Although Polyak's 1941 definition of 2.6 degrees radius is still commonly cited, I agree with the reviewer and have changed the text. I hesitate to use 'parafoveal' because using the same Polyak geography, 'parafoveal' means outside of 2.6 deg.

Line 412 – Fig 4 in this paper shows light-evoked elongation. The left panel has the only trace obtained from an individual/single cone. The right panel, on the other hand, is an average trace from 10 to 30 cones. It is very worthwhile educating the reader that these measurements are very challenging to achieve at the resolution of individual cones even with hardware adaptive optics and that averaging from multiple cones or multiple measurements are sometimes required to reveal these responses. This is a critical point and what makes the Zhang et al. 2019 PNAS study all the more striking since they measured light-evoked activity in a large population of cones, on a cellular scale, to separate them by spectral type.

Agree. A paragraph toward the end of the manuscript has been added, describing the tradeoffs between experiment duration, spatial resolution, and sensitivity.

Line 432 and 410 – make pupil size consistent

Done.

Line 439 : Insert "bleach" as unit for 0.007% bleach.

Done.

Line 438 – 445: How did this rod elongation in humans compare with mouse and with the Lu et al. study ? Also, how did the rod elongation amplitude and velocity compare with that observed in cones ? Did the response saturate in cones similarly as in rods? Some discussion of the differences in rods and cones is essential here, and also with previous work on rods in humans and mouse.

This has partly been addressed in Fig. 4, and also addressed in this paragraph.

Lines 476, 477 : The effective A-scan is just one metric by which to judge speed and its relevance to ORGs, though it is highly incomplete. It must be stated that the maximum possible temporal resolution in FF-SSOCT is equivalent to the volume rate, equal to 167 Hz or 6 ms in Hillmann PNAS 2017. In Zhang et al., PNAS 2019 and in the Pandiyan et al. 2020 papers, ORGs were obtained at speeds as high as 3 kHz and 16 kHz respectively by analyzing consecutive B-scans, but with the loss of all spatial information. However, such a high temporal resolution is beyond the reach of FF-SS-OCT, since it does not allow this form of cross-sectional analysis. Overall it is valuable here to more precisely describe the impact of speed and the paradigm of point vs. line vs. full-field, as it pertains to ORG and not merely rely on effective A-scan rates to draw a comparison.

Because this is a primarily clinical journal I don't want to go into too much technical detail, but a paragraph summarizing these differences has been added in the last section.

Line 482 – Please update references 110 and 111 to their published versions (PMID : 33031739 and 32917686). Pandiyan et al. Science Advances 2020 measured the rate of elongation across a large bleach range and showed

that the rate of elongation saturates in cones at ~15 % bleach. It would be very worthwhile for the reader to be educated with a complete description of the rates of elongation and how they saturate in cones (405-422).

Line 486 – Kudos to the authors for rightly crediting this group for the introduction of the ORG terminology to describe these measurements. For accuracy, one should cite instead the reference where this terminology first appeared to describe the measurements, in the same conference, in the abstract from the same group titled: "Optoretinogram: stimulus-induced optical changes in photoreceptors observed with phase-resolved line-scan OCT." (https://iovs.arvojournals.org/article.aspx?articleid=2741885)

To the editor: I have tried to strike a balance between this comment and that of reviewer 1 who didn't like this attribution. Hopefully both criticisms have been addressed.

Line 515-518: The fast contraction is not caused by the osmotic rectification of the ERP. Please correct. And see PMID : 33031739 and 32917686 for more details. It should also be very clearly noted that Zhang et al. 2017 mouse study found that only about 10 % of the elongation was attributable to transducin. Lastly, inspecting the family of curves in PMID 32917686 shows that the time scales are shorter than 10 ms, and closer to 5 ms.

This was too hasty a description. I have put it in terms of membrane tension instead.

Line 534 : 5nm is perhaps more accurate.

Modified.

Line 547 : One can measure a bleach or light-induced response in single or a handful of cones – isn't that a sufficient criterion for being more sensitive than perimetry and ERG that have far coarser resolution.

I agree with the reviewer, but hesitate to put things too strongly at this early stage in ORG development. Also I strongly believe that these kinds of comparisons should be done.

Line 613 : same as above, "attributing 10 % of the disassociation. . . " This discussion on transducin might benefit from a dose of skepticism in that mouse rods are different than human rods, and that only a small fraction of elongation was explained by transducin. Same with Lines 632 – 635 on the speculation of GTP-dependent deactivation. This is very much along the lines of the first comment. For a review, I personally think some speculation is healthy, but with some necessary caveats and explicitly stating that the mechanisms currently remain largely unknown.

I agree. I have added a "dose of skepticism" at the end of this section.

Line 676 : "weighted by spectral sensitivity PSF and OPL changes" – very unclear what is meant here. Please rephrase

Agreed. Edited for clarity.

Line 678 : over an area ranging between 0.07 – 0.27 deg square. Importantly, it could be mentioned that this was without adaptive optics and is thus encouraging for clinical translation.

Agreed. Changed.

Line 761 : I counted 8 published papers since 2016. Please check.

Agreed. Changed.

Line 780: Minor point – is it essential to qualify reference 129 as a "conference" paper? It would seem reference #128 and #28 are conference papers as well. Consider making it consistent when referring to all of them.

Agreed. Changed.