Author’s response to reviews

Title: Association of Photopic and Mesopic Contrast Sensitivity in Older Drivers with Risk of Motor Vehicle Collision Using Naturalistic Driving Data

Authors:

Cynthia Owsley (owsley@uab.edu)
Thomas Swain (thomasswain@uabmc.edu)
Rong Liu (liur@uab.edu)
Gerald McGwin (gmcgwin@uabmc.edu)
MiYoung Kwon (miyoungkwon@uabmc.edu)

Version: 1 Date: 13 Nov 2019

Author’s response to reviews:

Response to Reviewers’ Comments

Thank you for providing feedback from the four reviewers. We appreciate these comments and have incorporated them in revising the manuscript. Our responses to each comment are under each comment.

Reviewer 1:

“This paper investigated the relationship between photopic and mesopic contrast sensitivity and driving risk among older drivers using SHRP 2 naturalistic driving study data. The data, methods, results, and conclusion are scientifically sound. The writing is in professional manner.”

Thank you.

“(1) There are some previous papers that also studied the older driver vision condition and driving risk, especially also using SHRP 2 data. Please include (at least some of) them in your literature review part.”

We agree that it is important to refer in the manuscript to the literature pertinent to the topic of the manuscript – vision impairment and driving. The vision impairment and driving literature with respect to visual risk factors for collision involvement is enormous. In the past 20 years we have summarized this literature in three separate publications, and all of these overviews of the literature are cited in the present manuscript (ref #4-6). In addition, we have specifically included
citations to publications on contrast sensitivity and collision risk, since contrast sensitivity as a risk factor for collision involvement is the focus of this study (ref #13-18). We have also included additional references on contrast sensitivity and driving performance (ref #19-26) and also the relationship between self-regulation (e.g., reductions in driving exposure) and contrast sensitivity (ref #54-58). The article published by the Virginia Tech group examining the relationship between contrast sensitivity and collision risk using SHRP2 data is now cited in the Introduction (ref #36). However, unlike our study, they did not use the data to fit a contrast sensitivity model which provides a more comprehensive summary of visibility for a broad variety of spatial stimuli.

“(2) Please be very clear which SHRP 2 variables did you use from SHRP 2 data. You can provide a list/table of variables, e.g. miles/hours, age, contrast sensitivity measures .... And which variables (e.g. CSF) are derived from those raw SHRP 2 variables, and how to derive.”

The SHRP2 variables used in our analyses are described in detail in the Method section. If a variable from SHRP2 is not mentioned, we did not use it. We focused on variables in our analyses that address the specific research questions we are asking. Since the text in the Method section is detailed, we would like to avoid the redundancy of including a Table. We are explicit in the Methods as to which contrast sensitivity variables were used from SHRP2. We also explain in detail how the contrast sensitivity data from SHRP2 was fit by the contrast sensitivity model.

“In methods part, please provide details of Poisson regression models (equations, setups, variables), and how “age-adjusted” was performed? I guess you need to consolidate some age groups, e.g. 95-99 group, as there are very few people in that group.”

Poisson regression was described correctly and concisely in the methods section. It is unnecessary to provide equations and variables since these items are clearly described in the methods section. Additionally, adjustment for variables is standard in statistical modeling; no further explanation is needed for anyone. We are unsure of what the review is referencing regarding the consolidation of age. Age in the SHRP2 database is a categorical variable; in our analyses, age was categorized as presented in Table 1.

“4) Results part. Table 1. race, ethnicity, and gender are not that important, as they are not in your model. Instead, you should look at the distribution of important variables (e.g. # crashes, # miles, the raw and derived contrast sensitivities variables, etc.), and the distribution across different age groups (to see if the pattern vary by age group).”

We respectfully disagree with the reviewer in that we do believe it is important to report race, ethnicity and gender. Including these variables facilitates the generalization of study results and allows one to compare and contrast results from other studies. Furthermore, research funded by the National Institutes of Health in the US, which this is, requires the reporting of these variables as part of study results. The number of crashes, number of miles driven, and other driver
variables are included in the third paragraph of the results and do not need to be repeated in a table. Table 2 summarizes all contrast sensitivity variables, so these are already included.

“5) Table 2. You'd better give more summary statistics, such as min, max, 5, 25, 50 (median), 95, 95 percentile, max. Here you can also provide the distribution of the raw SHRP 2 contrast sensitivity variables which you used to derived your variables. In a word, show details of the RELEVANT variables.”

The presentation of the summary statistics the reviewer has requested is uncommon and unnecessary. The mean and standard deviation for the contrast sensitivity variables derived from models is presented in Table 2. Additionally, it is irrelevant to present summary statistics of the raw SHRP2 contrast sensitivity variables since these were not used in any regression analyses. As explained in the introduction, a contrast sensitivity function provides a comprehensive summary of visibility of a range of spatial stimuli compared to the contrast sensitivities at single target sizes, which is why the SHRP2 raw contrast sensitivity variables were not used in the analysis.

“6) Table 3. all crash and at-fault crash are OK, but at-fault crashes are typically proportional to all crashes, so essentially you may not find a big difference. You can develop models to test on more vision related driving conditions, such as at nighttime/daytime, good weather/bad weather, etc. SHRP 2 event table have lots of such driving environmental variables. Use your domain knowledge to explore vision-relation bad driving conditions. In such as, you need to compare bad condition to a good condition. Moreover, you can study the risk of crash with severity level 1-4 (your current "all crash"), AS WELL AS crashes with severity level 1-3 (exclude level-4 low-risk tire strike curb).

You may find significant and meaningful findings when focusing of different driving environment scenarios and/or crash types. ("all crash" is just too broad, but you can still present though).”

The reviewer raises some interesting questions, however, they are beyond the scope of this paper and their answers are immaterial to the interpretation of the current analysis results.

“(7) When you present summary statistics, you can use figures/plots to help present the distributions (or scatter plots for 2-D relations).”

We appreciate the reviewer’s comment; however, the use of figures and plots is not necessary. All data are adequately summarized and presented either in the text or in the tables using the statistics shown (mean and standard deviation).

“(8) In short, please present more details on relevant variables, and explore more vision-related driving scenarios and different crash severity types.”
This comment seems to be a summary of #6 and #7 above from this reviewer. Please see the above responses.

Reviewer 2:

“Thank you for this interesting paper.”

Thank you.

“Crashes were defined as events where the SHRP2 participant's vehicle made contact with any object (vehicles, pedestrians, cyclists, animals, tree, buildings), at any speed, including non-premeditated departures from the roadway where at least one tire left the paved or intended travel surface of the road.” - how were roadside kerbs treated?”

Roadside curbs, or any roadside barrier, were crash objects when the vehicle made contact.

“Re: missing contrast sensitivity function data - was this because the stimuli were below the ability of the patients to see? Or lack of time or availability of equipment to do testing. If the former reason, it is interesting to know that results could have been even worse. How was data treated if vision was too poor to attempt the task? If the latter, was this an issue?”

Missing raw contrast sensitivity data occurred because the participant did not complete the test and/or it was not recorded. It is unknown why the data are missing as SHRP2 codebooks and datasets do not explain. The majority of missing CS data falls into the completely missing category (96% of missing photopic CS and 76.5% of missing mesopic CS). Out of those who did not have a CSF model, only 4% of missing for photopic and 23.5 of missing for mesopic was due to the participant’s inability to see the stimuli. This is possible because CS measures were taken for both eyes and used to construct binocular CSF. In order to construct the models, at least 3 CS data points were needed since the number of free parameters is two. Given the low number of participants with missing photopic CSF because they could not see the stimuli, the results are unlikely to be worse for photopic measures. Mesopic CSF associations could be worse due to the percent of participants who could not see any stimuli.

“Was there a relationship between mesopic contrast sensitivity functions and ambient light levels or time of day when incident occurred? Although stratification may not be possible, a graphical, hence descriptive, representation of time of day (and season as that affects light levels) would be useful from a symbolic data science viewpoint.”

We agree that the time of day and or ambient light level could modify the association; however, data were not provided on the date, time of day, or any meaningful measure of ambient lighting in which incidents occurred. In addition, we do not have access to the video data and cannot redact this data.
Also, was there any evidence of modification of driving behaviours for the group with poorer mesopic contrast sensitivity functions or poorer photopic contrast sensitivity functions, that could have affected results?

SHRP2 did not collect data on self-reported modifications in driving behaviors.

The authors suggest those who had poor photopic contrast sensitivity functions may have modified driving behaviour. Although it cannot be compared to previous behaviours, it would be useful to ascertain percentage of night vs day time driving. Similarly, for those with poor mesopic contrast sensitivity functions. Re the latter, by way of analogy people with poor peripheral visual fields are less aware of their loss than people with poor central visual fields. In the same way, people with mesopic loss may think it is the usual reduction in vision associated with lower light levels, rather than anything untoward. Would be useful if such an analysis could be conducted with existing data to test that hypothesised reason.

What we stated in the Discussion was that previous studies have found an association between impaired contrast sensitivity and modifications in driving behavior by older drivers. We do not know whether this is the case for the SHRP2 older drivers because self-reported driving modification data were not collected in SHRP2. We do not have all the trip files from SHRP2 so cannot discern the percentage of day vs. night driving.

A figure depicting the CSF and curve fitting process for at least one participant would be useful for the reader.

In the revised manuscript, we have included a figure illustrating the CSF parameters and curve obtained from a participant with normal vision.

The authors should comment on which parameters were considered for inclusion in the regression model, and why other measures were not included.

All parameters considered for inclusion in the models were the CSF derived parameters and age group. This is stated in the methods section. No other variables were considered since contrast sensitivity is not known to be confounded by or have effect modification by any other collected variables other than age. Tom, I thought we included variables that are relevant to the question being asked.

Reviewer #3:

This is a well written and straightforward account of the study. The SHRP data are a great resource, and the findings have direct relevance to public safety agency screening test choices.

Thank you.
“The authors could provide more description of the Optec 6500 test. What is the subject asked to do during this test? How do values obtained with this test compare to values obtained using other testing setups?”

More detail is now provided on the Optec 6500 in the Methods.

“Other studies generate contrast sensitivity functions using different stimuli, staircase procedures and more trials. Does this matter when considering how CSFs generated with this test might compare to CSFs created with those other testing methods?”

The Optec 6500 is a screening contrast sensitivity test based on the FACT test. It is not intended as a device for measuring contrast sensitivity for mechanistic studies in the laboratory. For the purposes of studying driver safety, it is appropriate since vision screening at driver licensing offices use screening devices, not laboratory protocols.

“It is not entirely clear to me what you mean when you say that you confirmed that the template-fit strategy yielded a reasonable fit to the CSF data across subjects and viewing conditions.”

We have revised the text to be clearer: “We confirmed that this template-fit approach yields a reasonable fit to the CSF data across subjects and viewing conditions by evaluating the goodness of the fit with R2 for each individual curve-fit. More than 80% of individual curve-fits showed a R2 value greater than 0.8 and more than 90% of individual fits showed a R2 value greater than 0.5. Participants with an R2 less than 0.5 were removed from the study.”

“Are there other papers that have specifically tested whether the area under the contract sensitivity function is predictive of driving outcomes?”

Not to our knowledge, and we follow the contrast sensitivity and driving literature closely. This was the motivation for the study. There is no previous work examining the relationship between AULCSF as a risk factor for motor vehicle collision involvement in older drivers.

“The authors could be more specific about why they think this might be a useful predictor that's superior to peak contrast sensitivity.”

As stated in the introduction, the major advantage of measuring the CSF, rather than contrast sensitivity for a single target size, is that it constitutes a comprehensive summary of visibility for a broad variety of spatial stimuli, not just peak contrast sensitivity. During driving, the driver encounters a visual world consisting of a broad range of spatial frequencies. Thus it seems appropriate to us to examine collision risk based on a variable based on a comprehensive summary of visibility for a broad range of spatial target sizes.

“It would be helpful to see a bit more about how the VTTI staff determine fault than the one sentence that's in the manuscript. It appears that fault could not be determined in some cases.”

Fault was not coded by VTTI staff if there was no observed evidence from the video data.
“Why not use near-crashes in the analysis as well, given that other studies have found to be a useful surrogate for crashes, and crashes being rare makes statistical inference more difficult?”

Near crashes are not true crashes and are an overestimate of what we are examining, collision risk.

“How choose quartiles to analyze the AUC data rather than a continuous analysis?”

Contrast sensitivity function derived variables, AUC and peak log sensitivity, were categorized to provide referent groups and meaning to these somewhat arbitrary continuous variables. It is easier for those unfamiliar with these complex visual constructs to know that those in the bottom quarter of the sample did (or did not) have an increased risk of crash. Additionally, these variables might not always follow a normal distribution and quartiles allow for these variables to be easily modeled while accounting for non-normality.

“One of the most exciting aspects of naturalistic recording methods is the ability to know something about how much the subjects drives, where to, in what situations etc. This submission could say more about these elements when considering the link between contrast sensitivity and crashes. For instance, in the Discussion the authors discuss the fact that it is known that reduced photopic contrast sensitivity results in drivers avoiding complicated situations and reduce their exposure. But there is very little discussion of the conditions in which crashes occurred, or how contrast sensitivity affected the driving space, traffic conditions and roadway illumination, or mileage of drivers in this study. For instance, the authors' hypothesis about mesopic tests being useful predictors for day and night crashes due to the wide variety of light conditions encountered both day and night could perhaps be explored, given that there is video of all of the crashes and presumably one could say something about the light levels in the preceding time period.”

We agree with the reviewer that naturalistic driving provides enormous amounts of data. While ideally all possible items, like lighting conditions, could be explored, we do not we do not have access to raw video data from the SHRP2 study.

“How did contrast sensitivity predict exposure in this group of subjects?”

The study’s aim was not to develop a predictive model for driving exposure based on contrast sensitivity. We used the driving exposure data in order to examine the relationship between contrast sensitivity and crash involvement in our Poisson regression models. At the reviewer’s
request, we did look at the association between driving exposure and contrast sensitivity (Spearman correlations) and here are the results: Photopic AULCSF (rs=0.06, p-value=0.1088), photopic peak sensitivity (rs=0.03, p-value=0.3365), and mesopic peak sensitivity (rs=0.06, p-value=0.0793) were not correlated with total miles driven; mesopic AULCSF was significantly correlated (rs=0.08, p-value=0.0149). But this latter correlation is also very weak.

Generally, my opinion is that the manuscript would be improved by more discussion of how contrast sensitivity affected other outcomes that can be observed in naturalistic recordings in this study. Perhaps the authors feel that the sample size is not large enough, effective analysis methods do not yet exist, or they plan to report on these relationships later?

The focus of this paper is on contrast sensitivity and collision involvement. We feel that a focused look at this issue is appropriate given its importance to the literature. From the standpoint of the National Highway Traffic Safety Administration in the US, collisions are the most important highway safety metric. The manuscript would be grandiose in scope if it included many other outcomes. Thus, our focus on collisions is appropriate in our opinion.

Second paragraph of Discussion: remove apostrophe from "Driver's".

Done.

Second to last paragraph of the Discussion, last sentence: I would advise against the use of "this study" as I found it initially confusing whether you meant the study being cited or the study that is the subject of your manuscript.

Thank you for pointing this out. We have clarified what study we are referring to.

Reviewer 4:

“In this study, the authors report (1) presence of a correlation between mesopic contrast sensitivity and crash rate among older drivers; and (2) absence of a correlation between photopic contrast sensitivity and crash rate. Neither conclusion is novel, as the authors admit; however, the large and prospective study design is more robust than previous studies that have been cross-sectional, retrospective, or smaller. Another unique advantage of this analysis is its use of a recently-described spatial contrast sensitivity function, which accounts for object size and is more relevant to real-world visual function than traditional contrast sensitivity measurements such as the Pelli-Robson chart. Overall, the paper makes a strong case for the utility of mesopic but not photopic contrast sensitivity in the prediction of future crash rate among older drivers”.

Thank you.

In response to the reviewer’s comment that finding an association between mesopic contrast sensitivity and crash rate is not novel: On a further detailed reading of the literature we believe
that the finding is in fact novel. The previous literature related to this general topic can be summarized as follows. We have also included discussion of this literature in the Discussion section of the manuscript.

Kimlin et al. (PMID 28564700) report an association between mesopic contrast sensitivity and on-road driving performance by older drivers, but not links to motor vehicle collision involvement. Black et al. (PMID 31378990) found in young drivers that correction of astigmatism with toric contact lenses improved on-road driving performance which was also linked to better mesopic vision, however, they did not link improved mesopic vision to a reduction in collision involvement, nor was the study on older drivers. Von Hebenstreit (PMID 6482301) reported that commercial drivers (bus and truck drivers) with “reduced twilight vision” were more frequently involved in collisions, however, the focus in this study is on commercial drivers and it is unclear how “twilight vision” is defined. Lachenmayr et al. (PMID 9531801) reported that drivers with worse mesopic vision were more likely to be involved in night-time collisions. Some of the drivers in the sample were older drivers, however the vast majority of drivers (2/3 of the sample) were ≤ 60 years old.

While these studies address topics relevant to our finding on older drivers and increased collision risk for those with reduced mesopic contrast sensitivity, they do not themselves report this finding. However, we have now added these studies and their citations to our Discussion in the manuscript since they are relevant to the topic of our study.

“This paper could be strengthened by additional analyses that exploit the full dataset of the SHRP2 study. As the authors describe at length in the Background section, ‘naturalistic’ driving studies are superior to police accident report analyses because of the wealth of data collected by sophisticated vehicle sensors. However, this study ultimately analyzed only two outcomes—crashes and at-fault crashes—which frankly could have been gleaned from more ‘traditional’ data sets. It would be interesting to see if contrast sensitivity correlates with other SHRP2 data points that characterize high-risk driving, such as near-crashes or roadway departures.”

As we have pointed out above, the focus of this paper is on contrast sensitivity and collision involvement. We feel that a focused look at this issue is appropriate given its importance to the literature. From the standpoint of the National Highway Traffic Safety Administration in the US, collisions are the most important highway safety metric. The manuscript would be grandiose in scope if it included many other outcomes. Thus, our focus on collisions is appropriate in our opinion.

The reviewer mentions that crashes and at-fault crashes could have been studied in the context of more traditional data sets. We assume the reviewer means data sets based on accident reports. This is indeed the case, and we have done this type of research. The novelty in the approach here is the use of crashes and at-fault crashes that are derived from naturalistic driving recordings, not from accident reports. In the Introduction we discuss the limitations of using accident reports for studying crashes, and how naturalistic driving data can be useful in obtaining a more realistic understanding of risk.