RCS ISPR comments of Study Design					
Comment topic	**Reviewer**	**Location: Page/Line in original document**	**Reviewer comment**	**Author response**	**Reviewer response to author action**
Comments and responses from initial review in Rows 35-173. Comments and responses from second and final review in Rows 6-31.					
Definition and calculation of shade	**AE**	Pg.1, Purpose	R1 and R2 identified a lack of clarity around the definition of the term "shade." As R2 pointed out, this lack of clarity is partially due to an apparent error in the definition of Global Site Factor (GSF) as provided by Roon et al. (2021). From my reading of the Delta-T Hemiview manual, it appears that the software first computes time series of clear-sky global solar radiation ("global" means direct plus diffuse) reaching the ground under the canopy for a specified period, then integrates the time series. I imagine that the computations follow an approach similar to that applied by Moore et al. (2005, Hydrol. Proc., Eq. A1). This quantity is then divided by the time-integrated global solar radiation at an open site to generate an index that ranges from 0 (complete shading) to 1 (no shading).	Agreed, thank you for the helpful suggestions. The definition of shade has been updated and explained accordingly.	
Definition and calculation of shade	**3**		R3 identified several occasions for which there is a lack of clarity about whether shade or the change in shade was the response variable in the predictive models.	Agreed, thanks for catching that. We have tried to make it clearer throughout the text that the primary response variable is change in effective shade (ΔES).	
Acquisition of the hemispherical photographs	**1**	Pg. 6, Study site layout	Because hemispheric photography in stream channels is difficult and time-consuming, it is important to conduct some prototype studies and refine the methods from the experience. The authors still don't have a plan for prototyping and testing their photography plan. I infer that none of them have actually taken hemispheric photography photos in a stream under a forest, and they don't appreciate all the problems that go with it. You might have a plan to take photos at set distances, but when you get there, there is a wood jam, or a deep pool, or low woody vegetation, or something else, and you have to set the camera somewhere else. How do you randomize that choice and avoid implicit bias in location? It takes a while to get the camera set up at each spot and pack it up safely for moving to the next spot. The whole process is slow. And then there is weather. If you have a partly cloudy sky, or too much cloud cover, or rain, your photos aren't going to work for estimating canopy cover. It rains or is very cloudy much of the time in western Washington. Photos are better in flat light of morning or evening. Consistently good photography days will be limited to the period from July 5th to mid-September. They investigators really need to do some trial runs and revise the plan based on what they find out. They also need to evaluate the 7.5 foot spacing relative to canopy gap widths and gap frequencies in channels. 7.5 feet is very close together for hemispheric photography, and I'm not sure that a small number of photographs taken at the density will provide a good sample of reach-scale canopy cover. I would also suggest that they scale the distances to the channel width, e.g. take a photo every channel width. This will scale the spacing to the size of the channel. Furthermore, five photographs is not a lot to capture the variability of canopy cover on a forested stream.	• We agree that hemispherical photography in stream channels is difficult and time-consuming, and so have attempted to balance that difficulty with effort (i.e., number of photos per treatment/plot). • We agree that a testing phase could be helpful. We will pass this recommendation along to the stakeholders. • The CMER program has used hemispherical photography methods for past and present projects, and employs staff who are familiar with hemispherical photography methods. We plan to hire qualified contractors to help implement this study, including professionals with hemispherical photography skills and experience. • We have added a protocol for dealing with in-stream obstructions in the revised study design. Thank you for catching that. • The hemispherical photos will be specific to individual treatment plots each measuring 375 feet along the stream channel. The number (5) and spacing (7.5 feet) of hemispherical photos for each treatment/plot will ensure that shade measurements (hemispherical camera viewshed) for a given plot will not be influenced by areas outside of the plot for the study period of interest (solar altitudes of 40° or greater from 1 June to 1 September [Figures 3a and 3b]). If we added more than 5 photos per plot, and/or spaced the photo points further than 7.5 feet apart, then the camera viewshed would be capturing sources of shade originating outside of the plot, so we could not directly relate change in effective shade to the riparian stand conditions or harvest treatments within a given plot (which is our primary objective).	
Acquisition of the hemispherical photographs	**2**	Pg. 6, Study site layout	I have reviewed the revised version of the Riparian Characteristics and Shade Response Experimental Research Study and the comment matrix and feel that the authors have adequately addressed most of the comments. I'm not providing a detailed specific review since the request was to assess whether concerns were satisfactorily addressed but can provide a more detailed review if requested. The rationale that was provided for some of my main concerns – i.e. the non-cut side masking, solar elevation angle filtering, and reduced range of orientations make sense given the specific objectives of the study to explicitly identify direct effects of timber harvest. While this will address the study objectives, I feel that the overall scientific benefits of the substantial study investment would be enhanced by a more detailed assessment of effective shade changes encompassing the full range of hemispherical and stream conditions. I therefore recommend that a more detailed treatment be considered during the analysis phase since the hemiphoto processing steps are simple and inexpensive relative to the experimental treatments and field data collection. This could be accomplished by following the study plan as written for CMER's needs and report, but incorporating a more complete analysis for a subsequent journal article. The more complete analysis could be completed both with and without non-cut side masking, additional stream cut side orientations, and perhaps a simple model-based analysis to estimate the potential effects of limiting the solar elevation angles for cases where there is a distinct vertical difference in canopy optical density. I am still concerned that the limited number of hemiphotos used in the analysis will be adequate to capture the variations in stream shade. In the comment matrix it was noted that a statistician with expertise in these types of studies was consulted which is somewhat reassuring. It would however be much more convincing if the specific quantitative details of the guidance were provided in the document (possibly as an appendix) or if this aspect of the study design were supported with specific data either from an existing or pilot study. It is therefore not possible to critically assess whether this aspect of the study design is adequate to address the study objectives.	We were not able to test the number of photos in a quantitative way due to lack of relevant data from an existing or pilot study. The study design originally called for 3 hemispherical photos per plot, but this number was increased to 5 photos per plot, based on the dialogue with a statistician familiar with hemiphoto methods in similar stream systems. This required increasing the original plot size to accommodate more photo points, i.e., to avoid capturing shade originating from sources outside the plots. We attempted to maximize the number of photos per plot while balancing concerns about the difficulty and complexity of hemiphoto collection (as noted by R1), as well as budget/logistical constraints associated with increasing plot size. Given that this study requires an extensive, 100% tree inventory within each plot, keeping the plots to a manageable size was important for logistical and budgetary purposes.	

Study design and analysis	**AE**	Pg. 22, Analysis	Most of the comments on these aspects were provided by R3. That reviewer is mainly satisfied with the authors' responses. However, R3 made some comments that the authors need to consider as the study proceeds and the data are analyzed; see highlighted rows in the spreadsheet that contains the responses. In addition to the responses by R3, I have a comment related to the normality of the shade values. In response to the reviewer's comment on this point, the authors stated that: The text states, "Shade values will not be normally distributed; however, the differences in shade values will be approximately normally distributed." The text has been changed such that a mean, not median, of the five shade measurements will be taken for each plot. Means are normally distributed and differences in means are normally distributed. I would note that, according to the central limit theorem, the sampling distribution of a mean approaches a normal distribution as the sample size increases. If the population being sampled is highly skewed, a sample size of five may be insufficient to reduce the non-normality of either an individual mean value of shade or the difference between two mean values to a level sufficient to meet the assumptions of a statistical technique.	Thank you for the information. We have modified the text accordingly.	
Acquisition of the hemispherical photographs	**1**		The proposal is written as if hemispheric photography is easy and quick. In my previous review, I made the point that hemispheric photography is a real pain in the neck, and most people are going to do something simpler, and that it is important to compare hemispheric photography results with simpler techniques like densiometer measurements. This would be a simple way to make the study more useful, but the authors also ignored this suggestion. It's OK that they don't want to run shade models, but other people are going to want to use their data to test shade models, so they should make sure to collect all the data needed for common shade models. It would be a lost opportunity if such data weren't collected.	Thanks, noted. We address the issue of densiometer measurements below.	
General comment	**1**		It's OK that they don't want to run shade models, but other people are going to want to use their data to test shade models, so they should make sure to collect all the data needed for common shade models. It would be a lost opportunity if such data weren't collected.	Thanks, noted. We address the issue of densiometer measurements below.	
General comment	**2**	Pg. 28, Site availability and sample size	L1254: Does this really mean to read ">" in this sentence? It seems that "<" would be more appropriate. Please also provide a justification for why this site constraint is included. In the response matrix it says that sites will not be excluded based on overstorey density thresholds so this addition seems in conflict with the response matrix.	We doubled checked all of the ">" and "<" symbols in the document and confirmed that they are all accurate. It is correct that sites will not be excluded based on overstory density. The related text has been modified to improve clarity.	
Definition and calculation of shade	**AE**	Pg. 20, Hemispherical photo post-processing and analysis	As an overarching comment, I recommend that the authors use the term "shade" in a generic sense and adopt the use of a more specific variable name to describe the index of shade that is actually used in the analysis. These terms should be clearly defined near the beginning of the proposal and used in a consistent manner throughout. Specific recommendations following from the comments summarized in part A1, above, are provided below. (a) I believe that GSF as described in the Hemiview manual is the relevant index to use in the calculation of shade rather than an index representing canopy cover or gap fraction. As recommended by R2, the authors should review the Hemiview manual to clarify their understanding of how GSF is calculated, and to ensure that the analyses and interpretations are consistent with the physical meaning of GSF.	Agreed. We have updated and clarified the definition of effective shade based on your recommendations.	
Definition and calculation of shade	**AE**	Pg. 21, Hemispherical photo post-processing and analysis	The authors will need to clarify the definition of the response variable for each analysis – i.e., whether it is the actual value of "shade" or the change in shade associated with the treatment.	Agreed. The text has been updated throughout to clarify that the response variable is change in effective shade (ΔES).	
Acquisition of the hemispherical photographs	**AE**		I agree with R1 and R2 that five photographs per plot may be insufficient to capture the range of variability along a stream reach. In addition, taking the photographs at a fixed distance from the bank would not capture across-stream variability, and could introduce bias, at least for "wider" streams. I do not necessarily agree with R1's recommendation to scale the longitudinal sampling distance to the stream width, especially for narrow streams and small sample sizes. I have the following two recommendations: (a) Ideally, a pilot study should be conducted to document the spatial variability of shade both along and across a reach as a basis for designing a sampling scheme to generate unbiased estimates that have sufficient precision for the study purposes. (b) The study should include sampling at multiple locations across the stream in addition to along the stream, at least for "wider" streams, as indicated by the analysis of results from a pilot study.	These are good recommendations. We will inform the stakeholders that the reviewers recommend conducting some sort of pilot study prior to study implementation, which would require an adjustment of the pre-approved budget and timeline. As R1 correctly noted, "hemispheric photography in stream channels is difficult and time-consuming". We chose to take our measurements at a fixed distance from the bank with the realization that taking hemiphotos in the middle of a large stream may not be logistically feasible (i.e., too deep to wade and level a camera). We expect that some streams included this this study will not be wadeable across their entire width, and wadeability will vary from site to site. Additionally, taking photos at a fixed distance will help provide consistency among study sites with different stream widths (i.e., help isolate shade response to treatments within a plot), which will likely encompass multiple different stream widths.	
Study design and analysis	**AE**		Study design and analysis: Most of the comments on these aspects were provided by R3. That reviewer is mainly satisfied with the authors' responses. However, R3 made some comments that the authors need to consider as the study is implemented and the data analysed; see highlighted rows in the spreadsheet (Rows 35, 94, 134, 148, 152, 173)	Thank you for the helpful comments and suggestions.	
Study design and analysis	**AE**	Pg. 22, Analysis	The authors will need to acknowledge that, due to the small sample size, the mean shade values and their differences will very likely not follow a normal distribution. The authors should ensure that their analysis is robust to deviations from normality in the response variable, if normality is indeed an underlying assumption. However, as R3 pointed out, the key assumption underlying linear mixed effects models is that the error terms are normally distributed. This assumption will need to be checked as part of the analysis and appropriate remedial measures implemented if the assumption is clearly violated.	In the past, we have analyzed densiometer or hemiphoto shade using a beta distribution in a GLMM with an interaction for period. We get similar answers to the analysis proposed here and that is the alternative plan if necessary.	

General comment	**AE**	Pg. 5, Study site layout	I recommend that the authors include measurements using a densiometer to allow an evaluation of the potential utility of a simpler and less costly methodology for shade evaluation in an operational context. These measurements could perhaps be made at a subset of the sample sites.	The stakeholders determined that hemiphotography methods would be used at an early step in the development of this study. Testing the utility of other methods or comparing different methods is outside the scope of this study. The hemiphoto methods were determined to be most suitable for the objectives of this study because they: (1) are considered less subjective and thus provide greater consistency across measurements and observers, (2) create a permanent record of observations that can be queried at any time in the future, and re-analyzed under multiple other scenarios (e.g., shorter or longer periods of the year), and (3) can be analyzed in a standardized manner across all study sites; that is, the analysis software can be programmed to analyze the hemiphotos for effective shade (not just canopy cover) for a common solar period (time of year and range of solar altitudes), stream azimuth, and latitude/longitude. The densiometer methods cannot be standardized or analyzed in this way, making it more difficult to isolate the treatment effects on change in effective shade, which is our primary study objective.	
General comment	**AE**		This study will involve a major investment of funds and human resources, and it makes sense to make additional measurements that will maximize the utility of the data that are collected. Therefore, while I appreciate that while testing shade models may be outside the scope of the proposed study, I recommend that the field crews collect the data that would allow shade models to be tested in follow-up analyses, as long as those additional measurements do not compromise the collection of the core data.	We will pass this recommendation on to the stakeholders who determined the scope and budget for this project. Including it in the study design is not within the author's discretion at this time.	
General comment	**AE**	Pg. 24, Analysis	I found it difficult to understand the logic underlying this assumption and the threshold chosen for "taking action," which I presume from the following sentence refers to inclusion of an interaction term in the analysis. The authors need to clarify the criteria for site selection to ensure that they are consistent with the intended analysis. It is unclear why this assumption is numbered "1" when it appears to be the only assumption listed.	Agreed. The text has been modified accordingly.	
General comment	**1**		The study authors have done substantial editorial work to make the text clearer and improve the organization, but they basically ignored many of the substantive comments and suggestions from the first review. I have reviewed literally hundreds of journal articles and grant proposals, and I don't know if I recall any authors ever ignoring substantive comments to this extent. It is thus hard to be motivated to provide additional substantive review of the revised proposal. This is a difficult position for a reviewer to be in, and I don't think I have ever written a summary review paragraph like this.	We sincerely appreciated and considered all of your comments, and regret this was not well-communicated in our previous correspondence. We hope our enclosed responses better communicate our limitations for incorporating some of these comments within our stakeholder-driven process.	
Definition and calculation of shade	**1**	Pg. 1, Purpose	I still think this can be a useful and informative study, but the proposal still has serious flaws. Here are my main suggestions/comments on the revised proposal. 1. The proposal focuses on measuring "stream shade" although it doesn't define what is meant by stream shade. The importance of the definition was stressed in the previous review, but the authors simply removed any definition. Previously, they had provided two differing definitions, and rather than resolve the differences, they apparently elected to ignore any distinctions among the various stream shade definitions in the literature. This point is important because it affects the validity of the methods and the inference. It's hard to review a proposal when it isn't clear what is the fundamental unit the proposal seeks to evaluate. I infer that what they mean by stream shade is "canopy cover assuming an east-west stream azimuth". If that is what they mean, they should say it up front.	These comments have been incorporated.	
Definition and calculation of shade	**1**	Pg. 20, Hemispherical photo post-processing and analysis	On page 33, the following de facto definition is provided (note, terminology has changed from "stream shade" to "effective shade"): "Following thresholding procedures, effective shade (%) will be calculated using the formula (1-GSF) × 100, where GSF (Global Site Factor) is the number of "open" pixels where the sun path crosses the sky during the period from 1 June to 1 September for solar altitudes 40° or greater (Roon et al. 2021)." Essentially, this is a summertime-integrated percent canopy cover. It is the same thing that is measured instantaneously with a canopy densiometer, but a lot more complicated.	Hemispherical photographs provide a permanent record of the geometry of sky visibility and obstruction. We will use Hemisfer software to compute effective shade for each photo, which can be calculated based on a particular solar path (time of day/year), stream azimuth (east-west), and latitude/longitude (central Washington). Thus, effective shade as defined in this study is related to, but not exactly the same as, canopy cover. We determined that hemiphotography methods were well-matched for our purposes because they will allow us to analyze the photos according to specific criteria of interest to our stakeholders, and standardize these criteria across study sites, which will reduce the influence of variability not associated with the harvest treatments. Hemiphotography methods also give us the option of re-analyzing the photos under different scenarios that may be of future interest (e.g., a two-week period in late July). Additionally, stakeholders may prefer hemiphotos over other methods possibly because they are viewed as being slightly less subjective and because photos can be re-analyzed in case of dispute. We have CMER published studies which show the correlation between densiometer measurements and hemispherical photography, so re-testing this relationship was not determined to be a priority for this study.	
General comment	**1**	Pg. 20, Hemispherical photo post-processing and analysis	On the next page, the de facto definition of shade is constrained to specific azimuth assumptions as follows: "Using the Hemisfer photo analysis software, hemispherical photos will be analyzed for the central latitude/longitude in Washington (47.3826, -120.4472) and for (1) east-west oriented streams with the treatment bank assigned to the south; and (2) north-south streams with the treatment bank assigned to the east. Note, for north-south orientations, an east-facing treatment bank was selected for purposes of consistency, but shade values are expected to be similar to a west-facing treatment bank." Taken together, this means the proposal should be called, "Riparian treatment effects on summer canopy cover for east-west and north-south streams in the center of Washington State." This would be a lot easier to understand.	This study design previously went through multiple rounds of stakeholder review and (for programmatic reasons) we cannot change the title, but we have incorporated this description in the text.	
Acquisition of the hemispherical photographs	**1**		2. The proposal is written as if hemispheric photography is easy and quick. In my previous review, I made the point that hemispheric photography is a real pain in the neck, and most people are going to do something simpler, and that it is important to compare hemispheric photography results with simpler techniques like densiometer measurements. This would be a simple way to make the study more useful, but the authors also ignored this suggestion.	The AMP has implemented several studies that incorporated shade measurement and our stakeholders (including PIs from those studies) have made the determination that this study will use hemispherical photography with some awareness of the problems and limitations. This is not an element of the study that we have authority to change at this point.	
Acquisition of the hemispherical photographs	**1**		3. Because hemispheric photography in stream channels is difficult and time-consuming, it is important to conduct some prototype studies and refine the methods from the experience. The authors still don't have a plan for prototyping and testing their photography plan. I infer that none of them have actually taken hemispheric photography photos in a stream under a forest, and they don't appreciate all the problems that go with it.	We have stakeholders/cooperators with this experience, and we have not yet identified a PI for this effort. If a pilot effort is required to get our field protocols in order, we will implement one.	

Acquisition of the hemispherical photographs	1	Pg. 6, Study site layout	You might have a plan to take photos at set distances, but when you get there, there is a wood jam, or a deep pool, or low woody vegetation, or something else, and you have to set the camera somewhere else.	We added text to address this point.	
Acquisition of the hemispherical photographs	1		How do you randomize that choice and avoid implicit bias in location? It takes a while to get the camera set up at each spot and pack it up safely for moving to the next spot. The whole process is slow. And then there is weather. If you have a partly cloudy sky, or too much cloud cover, or rain, your photos aren't going to work for estimating canopy cover. It rains or is very cloudy much of the time in western Washington. Photos are better in flat light of morning or evening. Consistently good photography days will be limited to the period from July 5th to mid-September.	CMER cooperators have been collecting hemispherical photos for some time and we are aware of these issues.	
Acquisition of the hemispherical photographs	1	Pg. 5, Study site layout	They investigators really need to do some trial runs and revise the plan based on what they find out. They also need to evaluate the 7.5 foot spacing relative to canopy gap widths and gap frequencies in channels. 7.5 feet is very close together for hemispheric photography, and I'm not sure that a small number of photographs taken at the density will provide a good sample of reach-scale canopy cover.	We have tried to do a better job of addressing this concern in the document.	
General comment	1		I would also suggest that they scale the distances to the channel width, e.g. take a photo every channel width. This will scale the spacing to the size of the channel. Furthermore, five photographs is not a lot to capture the variability of canopy cover on a forested stream. 4. It's OK that they don't want to run shade models, but other people are going to want to use their data to test shade models, so they should make sure to collect all the data needed for common shade models. It would be a lost opportunity if such data weren't collected. 5. The authors have done a good job explaining why they are normalizing the streams to specific azimuths – basically because their riparian treatments are occurring only on one side of the channel, not both, and they want to know how that treatment works in best-case conditions.	Thank you. We are glad to hear this issue is better clarified now.	
Comments and responses from initial review below. Highlighted rows indicate Reviewer 3 comments (in column F) referenced in cell D17					
Definition of shade	**All reviewers**		There is a lack of clarity and consistency in the use of the term "shade" and its quantification. AE recommendations: The authors should compute what R1 called "FAR" and what R2 called "effective shade" in addition to their proposed index based on pixel counting, which R2 has noted is actually a canopy density for a portion of the sky dome. R1 recommended using a densiometer to estimate canopy density in addition to determining shade indices from the hemispherical photographs. This is a worthwhile idea and should not require an inordinate amount of extra time in the field. If densiometer measurements are found to correlate with effective shade, that could be a useful tool for operational situations given the relative ease of making densiometer measurements compared to hemispherical photography.	Please see changes to text and related author responses throughout this matrix. Note that we are interested in change in shade in response to harvests and blocking by site.	R3: Line 548: the first analysis is labeled "Determine how treatments affect **shade**". Should 'shade' be changed to 'changes in shade over initial conditions' ?
Definition of shade	1	Pg. 5 Effective shade section. First and second paragraphs	It is fundamental to a good shade study that the authors clearly define shade and the metrics used to quantify it. The term "shade" has been used loosely by the stream ecology community, such that the term means fundamentally different things in different published works. Often in environmental science, we can't actually measure the characteristic of importance, so we measure approximations or correlates thereof. This is often true of shade studies, and a variety of measures of shade and shade correlates have been published. To these points, the authors start by presenting the most stringent and accurate definition at shade as follows: *In this study, effective shade is defined as the fraction of solar radiation flux blocked from reaching the stream surface over a specified period of the solar cycle.* Let's call this the fraction of available radiation (FAR). This is essentially the same definition used in the Oregon RipStream study as presented in Groom et al., 2001, FE&M as follows: *An effort was made to take photos when the sun would not be in the picture (e.g. overcast conditions, when the sun was below the topographic horizon) and under dry conditions. Shade values were calculated from the photographs using HemiView 2.1 software (Delta-T Devices, Cambridge, UK) as one minus the June 30 Global Site Factor (1-GSF). The GSF is the proportion of both direct and diffuse energy under a plant canopy relative to the available direct and diffuse energy for the given site's latitude/longitude.* However, in the next paragraph in the methods, the authors propose a fundamentally different working definition of shade: *Effective shade values will be calculated as the percentage of pixels occupied by canopy within the portion of the photograph where the sun path crosses the sky during the period from 1 June to 1 September for solar altitudes 40 degrees or greater.* Let's call this summer canopy cover by hemispheric photography (SCCHP). These definitions are not equivalent and lead to very different inferences about landscape controls over stream temperature. Due to the effects of stream azimuth and topographic shading, streams can have the same amount of canopy cover but different levels of FAR. FAR is what determines the radiative portions of the energy budget and thus the dominant controls of stream temperature. To illustrate the importance of the distinction between SCCHP and FAR, of two similar streams flow east-west, and all the canopy is on the south side of one stream and all on the north side of the other stream, the two streams will have very similar canopy cover but completely different levels of FAR and very different stream temperatures (see Garner et al. 2017). In order to calculate FAR, the hemispheric photography data must be processed through software that accounts for azimuth, latitude, and elevation. Without post-processing the photography data through such software, as was done in the RipStream study, it is not possible to calculate incident solar radiation or FAR fro a stream segment. Which shade and canopy measurements should be taken? The pros and cons of common measures of shade and canopy cover are worth considering in thinking about what variables to include in this study (this is not a complete list, FYI): FAR, light reduction, direct shading, hemispheric photography, and densiometer measurements. Specifics about each are described in lines 7-11 below.	Please see changes to text and related responses throughout this matrix. There is now only one description for how shade values will be derived. The description for photo post-processing has been expanded to address this and other comments.	
Definition of shade	2		The description of how the shade response metric will be derived from the hemispherical images will produce canopy density within a sector of the hemisphere corresponding to a specific time period and solar elevation angle range which is not equivalent to effective shade. Hemiphoto analysis software can be used to estimate hourly above and below canopy global shortwave radiation for the specified time period and solar elevation angels of interest and estimated effective shade determined simply as (1-below canopy/above - canopy shortwave radiation). It is also not clear why the untreated side of the streams will be masked prior to photo analysis since this is not necessary to assess the effect of hypothetical treatments on effective shade. A more appropriate approach that would be more reflective of actual conditions would be to simply estimate above and below-canopy global shortwave radiation and filter the data to exclude values for solar elevation angles below the specified threshold to reduce effects of neighboring treatments.	See text for shade description revisions. The untreated side of the stream needs to be masked to help isolate the effects of the treatments on shade. Masking the untreaded side will remove non-treatment effects and variability associated with the untreated side of the stream. For example, the untreated side is likely to vary among sites in terms of tree density, tree height, time since past logging occurred, etc. For the purposes of this study, actual shade values are less important than change in shade values.	
Definition of shade	2	Pg. 5, second paragraph and top of pg. 17	As noted above, the definition of effective shade in the 2nd paragraph at the top of pg. 5 and likewise on the top of pg. 17 is incorrect.	This topic has been refined in the text based on the collective comments.	

Definition of shade	**3**		Effective shade is actually not shade but the proportion of sky occupied by canopy. I wondered if the scientific community would all agree that canopy cover could be labeled shade. I note this only because I wondered about it and also wondered if it would be more accurate to rename your response to more accurately reflect what it is and then discuss why the response is a shade metric. But I'm not the expert in shade so leave that up to the authors.	Refinements have been made to the text accordingly.	Thank you
Research approach	**3**	Pg. 2-4	I appreciated the overview of the investigation on pages 2-4 prior to the more detailed explanation in the rest of the document. But some important points are missing in the section which led me to be confused as I continued to read. I suggest including a more explicit description of the temporal progression of the no-harvest-zone-width levels and the timing of the data collection (photographs) for each treatment combination in the beginning. A more explicit description of the chronological implementation of the harvest treatments would make the description of the process more clear. Specific edits suggested by reviewer were added to the study design doc in track changes by PM.	Agreed. Specific edits suggested by reviewer were incorporated. The figure (Figure 4) has been updated to more clearly show the harvest and photo sequence.	Thank you - the revision is MUCH more clear.
Research approach	**3**	Pg. 2 first paragraph under research approach section	Study is identified as a 'split plot' design with 2 factors and the levels of the factors are identified. The paragraph refers to Table 1. Further in the document (Analysis section), it is correctly noted that the design is not technically a split plot design since the split plot factor (zone width) is not randomly assigned to each whole plot (the plot that receives a thinning intensity). Zone width is assigned in the same order every whole plot. When the 'split plot factor' is applied in the same order in all whole plots, the design is not a split plot design but is a 'strip plot design'. Most experimental design texts will have this identified. However, this design is not exactly this either given the imposition of the ecoregion factor. I have included more details when reviewing the Analysis section.	See text and author responses throughout for changes to split plot language and descriptions.	Thanks - changes are helpful.
Research approach	**3**	Pg. 3 Table 1	Table 1 incorrectly identifies the 100' zone width as only existing in the plots assigned to the clear-cut thinning intensity. This is not correct. As shown in Figure 1, the 100' width occurs in every plot assigned to a thinning intensity. This particular level of width only exists when no thinning intensity has been implemented within the 100' strip of land so that may need clarification in the text as well. But technically the 100' width occurs in every "whole plot" that receives a thinning intensity. This could be interpreted as the 100' width being applied not to a plot but to the entire site. The LMM analysis could reflect that (replication for the 100' width is achieved through sites, not plots) but it's a complication to the analysis and likely requires a statistician's involvement in order to be that contrasts and standard errors are correctly estimated. An alternative approach could be: (a) be sure that photos are taken in in all 3 plots after Step 1 (see later comment about describing timing of photos) but before Step 2 and (b) Be sure that measurements for the 100' width are made independently in each plot at each site. Then treat the study as though there are 12 treatment combinations – each level of zone width with each level of thinning intensity. What will be true is that 3 of the treatment combinations (100' with moderate thinning, 100' with heavy thinning and 100' with clear-cut) are all equivalent. The analysis section of the manuscript notes that they can be averaged to estimate the effect of the 100' width. Then the analysis could proceed as a fully factorial design with all 12 treatment combinations with the understanding that 3 of them are identical. The advantage of having the 100' in each plot is that comparisons involving the 100' width and other widths will use the 100' width value from the specific plot which may be less variable than if the comparison involved an average over the 3 plots.	The table has been revised to indicate that the 100 ft treatment will be analyzed separately. Otherwise, no change needed because what Reviewer 3 describes regarding the photo sequence matches what is being proposed. The table depicts the different riparian *treatment level combinations* that will be included in this study, not individual plot configurations (the table does not specifically depict Plots 1, 2, and 3). Figure 4 shows how the treatments will be sequentially applied within each plot. Figure 4 has been expanded to clarify that photos will be taken after each treatment in each plot.	Thanks, the revisions clarify this.
Research approach	**3**	Research approach section	In this section of the manuscript, the timing of the photographs (when they are taken) is not described. I suggest including a more explicit description of the temporal progression of the no-harvest-zone-width levels and the timing of the data collection (photographs) for each treatment combination on page 3 in the Harvest sequence section. It is never explicitly described here and I was left fairly confused as I read through the rest of the document.	Comments incorporated. Figure 4 (formerly Figure 1) has been expanded to clarify that photos will be taken after each treatment in each plot.	Thank you - the revision is MUCH more clear.
Research approach	**3**	Pg. 3 Table 1	Please include a description of the timing of the photographs in the caption. Figure 1. I found the phrase "three harvest sequences" in the caption to be confusing – there are 4 rows in your figure but you say '3 sequences' and you want the reader to read 'down' the rows to see the sequences. One suggestion is to say 'Site layout and the progression from 100 foot no-harvest width (1) to 25-foot no-harvest-width (4) for 3 plots along one stream in this study. A compass rose could be helpful since the legend says that the figure is for an east/west orientation, but I think what is meant is that the top of the page is north.	Comments incorporated. Figure 4 (formerly Figure 1) has been expanded to clarify that photos will be taken after each treatment in each plot.	Thank you - the revision is MUCH more clear.
Shade measures	**1**	Pg. 5, Effective Shade section, first paragraph	Photo density in each segment length? I.e. separation between photos? (comment made within study design doc)	This information is provided in the Site Layout section.	
Shade measures	**1**	Pg. 5, Effective Shade section, third paragraph	I couldn't make heads or tails of this. Why not treat azimuth as a continuous variable and account for it? You can then use the data to test and evaluate shade models that can then be used to extrapolate findings to the spectrum of stream azimuth conditions (and other channel conditions). (comment from within study design doc)	See text for expanded explanation. Stream azimuth will be standardized to east-west and north-south streams using Hemisfer photo analysis software, regardless of actual stream azimuth in the field. This will provide a way of standardizing the stream azimuth variable across study sites. Treating azimuth as a continuous variable is outside of the scope of this study. We are specifically interested in what happens to shade response at north-south and east-west stream orientations - the range where maximum effects on shade are expected to occur.	
Shade measures	**1**		The pros and cons of including FAR (defined in line 4 above) in this study should be considered. This requires regularly spaced hemispheric photographs with careful alignment of the camera at each photo spot and requires post-processing with software that calculates incoming radiation based on the orientation of the canopy cover with respect to solar position (a function of date and latitude). FAR is difficult and expensive to measure, and it has rarely been done except in a few research studies.	See text and responses throughout this matrix.	

Shade measures	**1**		The pros and cons of including Light Reduction in this study should be considered. This requires measuring incident light at the stream level and comparing it to incident light measured in the open at the same location. This is essentially a direct measurement of FAR (above). The drawback is that you need a lot of light sensors and the light sensors must be kept level and free of litter fall during the study.	See related responses throughout this matrix. This suggestion seems to deviate from the main objectives and logistics of this study design.	
Shade measures	**1**		The pros and cons of including direct shading in this study should be considered. Incident solar radiation comes in two forms, direct and scattered. Shading of direct light by the canopy can be observed visually by noting what is in shadow and what is in direct light. Measurement of direct shade is fairly simple. See Li et al. 2012 (citation provided in submitted ISPR comments). The problem with direct shade is it does not provide a direct measure of scattered radiation.	No change needed. Multiple methods are available for studying light conditions in forested environments. The hemiphoto method was approved by committee for this study during the scoping phase. We are most interested in change in shade in response to the treatment level combinations.	
Shade measures	**1**		FAR is difficult to determine but provides much more insight into stream temperature variability and the landscape controls thereof. If the authors are going to the trouble to take hemispheric photographs of each stream, they should calculate and analyze the more informative FAR as well as the simpler SCCHP and even estimate the much simpler and faster canopy cover as measured by a densiometer. The paper below illustrates the importance of azimuth, latitude, and bank height on stream shade, and why these variables need to be accounted for in stream temperature studies. Li G., C.R. Jackson, and K.A. Kraseski. 2012. Modeled riparian stream shading: Agreement with field measurements and sensitivity to riparian conditions. Journal of Hydrology 428:142-151. The findings from this paper raise additional questions about the experimental design: 1) Why not calculate cover and FAR based on actual azimuth, rather than an artificial forcing of E-W and N-S orientations? You can also apply artificial azimuth values and run the shade and FAR calculation software for a spectrum of azimuth values. For example, see Garner, G., Malcolm, I.A., Sadler, J.P. and Hannah, D.M., 2017. The role of riparian vegetation density, channel orientation and water velocity in determining river temperature dynamics. *Journal of Hydrology*, *553*, pp.471-485. 2) Why not measure temperature as well? If you are going to the effort to conduct a large number of riparian/forest manipulations around the state, why not collect data on how stream temperatures also responded to these manipulations? Temperature loggers are relatively inexpensive and are very easy to use. The cost of adding a temperature study is small relative to the other study costs. 3) Why not use the collected canopy cover and FAR data to test existing models that calculate shade? There are a number of these models, and if they are shown to work over a robust dataset like the one that can be created here, then these models can eliminate the need to conduct this type of study in the future.	No change needed - see explanations throughout matrix. The hemiphoto method was selected by the group during the scoping phase of this study. Densiometer measurements were not included because they are more prone to oberserver subjectivity, and we do not have the staffing capacity/budget for conducting these analyses/comparisons. Assigning E-W and N-S orientations is a critical component of this study (see Reviewer 2 comment below). This will standardize shade response to the harvest treatments across sites, regardless of actual stream orientations in the field, providing comparability among sites. See further explanation in text and below. Measuring stream temperature is outside the scope (and budget and timeline) of this study. Most importantly, given that the entire sequence of harvest treatments will be applied within a short period of time (~10 days), we would only be able to measure temperature response to a given treatment over 1-2 days, which would not be meaningful.	
Shade measures	**1**	Pg. 4, Figure 1 caption	How will you account for differences in shade and temperature sensitivity to riparian cover due to azimuth and latitude? (Comment from within study design doc)	Latitude and stream azimuth will be standardized to east-west and north-south streams using Hemisfer photo analysis software, regardless of actual stream azimuth and location in the field.	
Shade measures	**2**	Pg. 4, Figure 1	For clarity, add a baseline (or step 0) case of a fully forested condition.	Figure 1 (now Figure 4) has been revised accordingly.	
Shade measures	**2**	Pg. 4, Figure 1	For clarity, add 25, 50, 75, and 100 foot markings on the sides of the figure.	Figure 1 (now Figure 4) has been revised accordingly.	
Shade measures	**2**		The material in the various sections about potentially using hemiphotos for other time periods of interest - specifically leaf-off periods - should probably be removed or modified. This limitation is noted later in the document, and again it is recommended that use of the images to assess other time periods where sun angles are much lower not be recommended given the potentially confounding effects of neighboring units and potentially large amount of data that would be filtered if higher solar elevation angle thresholds are used.	Agreed. Language related to the leaf-off period has been removed.	
Shade measures	**3**	Pg. 5, first paragraph	Paragraph 1 was confusing and lacks clarity. I suggest integrating the information from the more clear description beginning on page 13. The page 2 description inaccurately says that shade is estimated for each treatment – it's really estimated for each plot. The first paragraph says that there will be 5 photo collection intervals and refers to Figure 1. But Figure 1 shows the 5 photo locations, not intervals. The paragraph should describe the start, end and length of the intervals.	Figure 1 (now Figure 4) has been expanded for clarity. There are in fact 5 photo intervals and 5 photo points per plot, so no change is needed. Shade will in fact be estimated for each of the 10 treatment level combinations.	Thanks. Having the figure include the intervals helps.
Shade measures	**3**	Pg. 5, third paragraph	I'm not an expert in photo analysis so my comments may reflect my ignorance. I did not understand this paragraph at all when I first read it and also after reading pages 16- 17; so, I think that there is likely not enough detail/context. The first sentence "The same set of photos will be used to estimate effective shade for streams with (1) east-west and (2) north-south orientations, regardless of actual stream orientation in the field" could be more clear by explaining the process rather than the outcome – i.e., "Regardless of actual stream orientation in the field, in the software, the stream will be reoriented to lie in an east-west and a north-south orientation". But I am still puzzled by this notion of 'reorientation' and why this is needed. I think that more description of why this is an appropriate technique is needed for non-experts. Here is my (possibly erroneous) thinking. The path of sun across the stream, and relative to the stream, in situ, will change depending on the actual stream orientation. The stream-adjacent vegetation, which may then be present in pixels within the photo, is also likely a function of the actual stream orientation. So, the measurement (proportion of canopied pixels) is a function of the actual stream orientation even if you change the virtual orientation within the software. That is, virtually reorienting the stream direction, redefining the solar path based on the reorientation will indeed give you a different value for your shade response, but it is STILL a biological function of how the plants around the stream grew based on the actual orientation in the field. So, I really don't understand how the virtual reorientation will give you a measurement of shade that is biologically defensible. As I said, I may be all wrong here so will defer to the experts but given my confusion, some clarifications are	See text for expanded explanation and updated figure.	The expanded explanation is helpful.
Canopy Cover measures	**1**		Hemispheric photography should be included in the study because it provides the most accurate estimate of canopy cover from the camera elevation to the zenith and allows FAR calculation through appropriate software. Taking the photographs is slow and difficult. Camera must be leveled and aligned with north and set low to the ground (channel) to pick up the effects of shrubbery and herbaceous vegetation near the channel. Needs weather conditions with good and relatively consistent sky contrast from the canopy. If weather conditions aren't favorable, the day is shot. Photos must be repeated longitudinally to capture the patchy nature of canopy and canopy openings.	Agreed, thanks. There will be a detailed SOP written for hemiphoto collection following finalization of this study design.	

Canopy Cover measures	1		Densiometer measurements should be included in the study. Estimating canopy cover with a densiometer is easy and quick. Consequently, this tool has been used in many studies of riparian canopy cover, and it is the method used operationally. However, there are few studies that compare easy-to-collect densiometer measurements with more rigorous and densiometer measurements compare to FAR (and stream temperature change) would be very useful. It would be easy to add densiometer measurements to the study, and this would allow comparison of FAR, SCCHP, and densiometer measurements across a range of forest types and stream characteristics. In practice, people are far more likely to use a densiometer to measure SCCHP. See: Ringold, P.L., Van Sickle, J., Rasar, K. and Schacher, J., 2003. USE OF HEMISPHERIC IMAGERY FOR ESTIMATING STREAM SOLAR EXPOSURE. JAWRA Journal of the American Water Resources Association, 39(6), pp.1373-1384.	No change needed. It was a group decision to use hemiphoto methods, and to exclude densiometer measurements from this study. Comparing densiometer methods to hemiphoto methods is outside the scope of this study.	
Re-orienting photographs to simulate different stream azimuths	1		Why not calculate cover and FAR based on actual azimuth, rather than an artificial forcing of E-W and N-S orientations? You can also apply artificial azimuth values and run the shade and FAR calculation software for a spectrum of azimuth values. For example, see Garner, G., Malcom, I.A., Sadler, J.P. and Hannah, D.M., 2017. The role of riparian vegetation density, channel orientation and water velocity in determining river temperature dynamics. Journal of Hydrology, 553, pp. 471-485	No change needed. Assigning E-W and N-S orientations is a critical component of this study (see Reviewer 2 comment in row 28 below). This will standardize shade response to the harvest treatments across sites, regardless of actual stream orientations in the field, providing comparability among sites. Understanding shade response to orientations other than E-W and N-S is outside the scope of this study. The photos collected in this study will be available for future analyses of orientations other than E-W and N-S.	
Re-orienting photographs to simulate different stream azimuths	2		The idea of re-orienting hemispherical photographs to assess how shade loss varies with stream orientation is an excellent idea – assuming that there are not significant differences in the canopy conditions on opposing aspects. The component of the project that deals with assessment of aspect sensitivity is poorly developed in terms of the objectives and is largely lacking from the data analysis section. Specifically, the project will assess shade changes across just southside and eastside/westside harvests, and it's not clear why northside harvests are excluded, even though the shade changes will probably only affect the diffuse shortwave radiation component based on the data filtering on solar elevation angle.	First point - Agreed, thanks. Second point - Assessing shade for northside harvests is outside the scope of this study. This study focuses on the range of stream orientations that are most likely to have the largest effects on stream shade (based on solar geometry). Based on solar geometry, northside harvests are expected to have the smallest effect on stream shade. Because this study is being completed within the context of forestry regulations, it is important to focus limited resources on studying the range of azimuiths where the greatest effects due to riparian harvest are expected to occur. An expanded explanation has been added to the text.	
Re-orienting photographs to simulate different stream azimuths	2		The sensitivity of shade change to stream orientation should be more fully developed – either to assess how changes in shade are expected to vary either for end-member (e.g. north and south-side harvests) and median (east/west-side harvests) cases, or for the full range of stream orientations (e.g. in increments of 45° or less) since this could easily be done by re-orienting the images to gain a better understanding of potential shade changes across the broader population of stream reaches. This component needs to be improved both in the objectives/critical questions and explicitly in the data analysis procedures. There is also the potential to estimate the effects of 2-sided harvests by digitally halving and mirroring the treatment portion of the images. The effect would be expected to be minor since only the diffuse radiation component would be affected due to the low solar angle data filtering but could nonetheless be informative to better understand the full range of potential harvest impacts.	No change needed. This study targets the range of stream orientations where riparian harvests are expected to have the greatest potential effect on shade based on solar geometry. This information is important to understand from a forestry regulations/environmental standards/adaptive management context.Evaluating the diffuse radiation component is outside the scope of this study. Using a 2-sided mirror image approach is not necessary because we are estimating shade (%) based on the number of "open" pixels in the photo.	
Re-orienting photographs to simulate different stream azimuths	2		As noted above, the idea of re-orienting hemispherical photographs to understand how stream orientation may affect shade loss is a good approach, but this necessarily assumes that there are not significant differences in the canopy conditions on opposing aspects. Distinct aspect differences are probably unlikely in the wetter forest types and for north-south trending streams, but this assumption could be violated if there are east-west trending streams in the drier forests, so it is recommended that site homogeneity be carefully evaluated prior to using images to assess orientation differences.	Agreed. The stand inventory data will provide detailed information about stand composition and structure in each plot/treatment, so these variables will be considered in addition to stream orientation. Stand inventory data will also provide information about the variability in these characteristics among sites.	
Re-orienting photographs to simulate different stream azimuths	2		The approach of using hemiphotos to artificially explore variations related to stream orientations by specifying different photo orientations is an excellent idea. It would be good to either note the assumption that there are negligible differences in canopies on different stream orientations, or if possible, provide a reference to note how much canopy variation is expected to be associated with different stream aspects. This is probably not a concern in wetter systems, but there can be considerable differences in canopy conditions within 100 ft., between north and south-facing slopes in drier forests.	Agreed, thank you.	
Re-orienting photographs to simulate different stream azimuths	3		I am still puzzled by this notion of 'reorientation' and why this is needed. I think that more description of why this is an appropriate technique is needed for nonexperts. Here is my (possibly erroneous) thinking. The path of sun across the stream, and relative to the stream, in situ, will change depending on the actual stream orientation. The stream-adjacent vegetation, which may then be present in pixels within the photo, is also likely a function of the actual stream orientation. So, the measurement (proportion of canopied pixels) is a function of the actual stream orientation even if you change the virtual orientation within the software. That is, virtually reorienting the stream direction, redefining the solar path based on the reorientation will indeed give you a different value for your shade response, but it is STILL a biological function of how the plants around the stream grew based on the actual orientation in the field. So, I really don't understand how the virtual reorientation will give you a measurement of shade that is biologically defensible.	See expanded explanation in the text and responses throughout this matrix. The stand inventory data will help us explain differences in stand composition and structure across sites.	It's much clearer now, thanks.

Re-orienting photographs to simulate different stream azimuths	**3**	Pg. 22, end of the 2nd full paragraph	On p. 22 at the end of the 2nd full paragraph it says that the re-orientation of the streams to standard orientations (N/S, E/W) "will ensure that shade response to the treatments is not influenced by differences in stream orientation across sites". I've noted this earlier, but it comes up here too - I question whether this is biologically reasonable. The under and over-story growth on a site is a function of the site and environmental conditions at a site and that includes the orientation of the stream relative to the tree stands. Re-orienting the image doesn't change the *in situ* conditions that affect plant growth and vigor. So, while some conditions may be standardized by the re-orientation, others are not. I suggest you clarify what conditions can and cannot be standardized.	Stream orientation will be standardized *in terms of solar geometry - the solar path for the specified study period* . We acknowledge that vegetation conditions will not be standardized across sites and stream orientations. The 100% stand inventory data will provide detailed information stand composition and structure in each plot/treatment, so these variables will be known. Understory vegetation influences will be greatly reduced in this study due to the solar altitudes being analyzed (40° or greater) and the elevation of the camera (1 m).	Ok. Thanks for the explanation.
Re-orienting photographs to simulate different stream azimuths	**3**	Pg. 16/17, data analysis procedures, effective shade section	In other comments I tried to explain some of my confusion with the re-orientation of photos. I suggest consolidating information on Page 5 with this section together so a reader can obtain the whole picture at one time with consistent language.	Agreed, the document has been substantially restructured accordingly.	Thanks, it's much better
Re-orienting photographs to simulate different stream azimuths	**3**		I suggest adding additional language to explain how (or if) the effective shade measurements change under reorientation.	An expanded explanation has been added to the text.	Thank you.
Re-orienting photographs to simulate different stream azimuths	**3**	Pg. 5	It is not clear to me how reorientation of the stream in the software "*…will standardize estimates of effective shade according to stream orientation*". Effective shade is the "percentage of pixels occupied by canopy within the portion of the photograph where the sun path crosses the sky during the period from 1 June to 1 September for solar altitudes 40° or greater" (p 5). My understanding is that the 'sun path' depends on the true cardinal directions at the site. The text reads as if the reorientation will redefine the sun path so that a different value for effective shade will be determined under different orientations of the photo. Otherwise, why bother to reorient?	"The text reads as if the reorientation will redefine the sun path so that a different value for effective shade will be determined under different orientations of the photo." - This is a feasible assumption. This comment is in alignment with what we are proposing.	If this is an assumption that is implied it should probably be added to the text
Re-orienting photographs to simulate different stream azimuths	**3**		Another suggestion is to give an example photo and identify the sun path, the canopy and an effective shade measurement. The figure on page 17 isn't too helpful since it simply shows the direction of the stream being changed but the sun's path is not present.	The figure has been updated to improve clarity.	Thanks!
Re-orienting photographs to simulate different stream azimuths	**3**		If the idea is that the sun's path changes in the photo with reorientation, some explanation about why this is biologically meaningful will be needed. See previous comments.	See responses throughout. Relevant changes have been made to the text.	The changes helped, thanks.
Re-orienting photographs to simulate different stream azimuths	**AE**		Based on the comment by reviewer 1, it is not entirely clear in the proposal whether shade would be calculated using both the actual stream azimuths and the simulated E-W and N-S orientations, or just the simulated orientations. This point needs to be clarified. I would also encourage the authors to consider the proposal to compute shade for the actual stream orientations and four simulated orientations (N-S and E-W, as proposed, and also NW-SE and SW-NE). The authors need to develop the rationale for the re-orientation analysis more clearly in order to generate an approach that will be meaningful in an operational situation and is biologically defensible. As reviewer 2 points out, it may be that aspect-related contrasts in shade response may not be substantial in some conditions but could be in others. This point will need to be considered as a potential source of error in the final results, at least qualitatively.	See responses throughout and expanded explanation in the text.	
Taking hemispherical photographs in the field and post-processing the images	**1**		Experience is very helpful. I'm concerned about the time and manpower it will take to get the photographs taken for this many sites and manipulations over such a broad geographic area. How many teams will be taking the photos? If more than one team, they need to be trained together and some comparisons done of cover and shade measured by the two groups on different days (when they aren't together). I would recommend that the authors conduct a feasibility study to see how quickly and reliably the photo surveys can be done.	Agreed that this will be a complicated study to implement. These complexities have been communicated with the oversight committee, and informal feasibility investigations are currently underway. It is anticipated that this project will take place over multiple years.	
Taking hemispherical photographs in the field and post-processing the images	**2**		It would also be good to note how exposure settings for the hemispherical images will be determined. If the camera has a high dynamic range (HDR) setting it may be possible to automatically collect a series of images with different exposures to select the best image during the post-processing phase.	Agreed. Comment has been incorporated into the study design. Multiple exposures will be used during hemiphoto collection.	
Taking hemispherical photographs in the field and post-processing the images	**2**		The study design would benefit from some details on what software package will be used and how photos will be collected in the field and processed. Suggestion: If possible, it would also be good to estimate cover at both ~1 m and the water surface with a hemispherical densitometer for a more quantitative estimate of shade generated by low growing and understory vegetation and for consistency with other studies. As noted above, it is not clear why the untreated side of the streams will be masked from the analysis, and this is not necessary to assess the effect of hypothetical treatments on shade. A more appropriate approach that would be more reflective of actual conditions would be to simply estimate above and below-canopy global shortwave radiation and filter the data to exclude values for solar elevation angles below the specified threshold to reduce effects of neighboring treatments.	Hemisfer software details have been added to the text. After several discussions and iterations, it was a group decision to exclusively use hemiphotography methods at the height of 1 m. The dependent variable in this study is *change* in shade in relation to the sequential harvest treatments, with blocking by site. Our hemiphoto collection set up will help reduce the influence of low growing vegetation, helping to isolate the overstory treatment effects on shade.	
Taking hemispherical photographs in the field and post-processing the	**2**		One other suggestion for QA/QC is to do an assessment or duplicate effective shade analyses with 2 separate operators, since there is some subjectivity both in selecting the images with the best exposure and in thresholding the image to discretize sky from canopy in the images.	This suggestion has been added to the QA/QC section.	
Taking hemispherical photographs in the field and post-processing the images	**3**	Pg. 14, second paragraph	One bullet point says, "Hemispherical photographs will only be collected under uniform sky conditions when the sun is not directly in view, and according to the camera manual" and there is a somewhat different sentence on page 14 that says, "Photographs will be taken when no direct sunlight is visible, at pre-dawn, post-sunset, or under an evenly overcast sky." I suggest you decide which is most specific and use common language for both. There are conditions under which the sentence could contradict each other.	Deleted redundant information in QA/QC section and retained more detailed description earlier in document.	Thanks.

Study design - site selection and study population	**3**	Pg. 10-11, Figure 4a and 4b	I don't fully understand how the hemispherical photos 'work' so my comments may be irrelevant but want to provide them to illustrate how a reader may be confused. The figure shows shadow length of a tree located at the photo-taking location near the plot. I thought I understood that the photos would be of the canopy over the stream. Is the reader supposed to infer that a tree located along one of the shadow lengths in the figure (not at the photo point) would be captured by the photo taken along the stream? It would help to explain why imagining a tree located at the photo-taking point is relevant.	No change needed. "Is the reader supposed to infer that a tree located along one of the shadow lengths in the figure (not at the photo point) would be captured by the photo taken along the stream?" - Yes. The figure does not ask the reader to imagine a tree located at the photo-taking point. The 'imaginary tree' is at the outer extent of the lines relative to the central photo point.	Thanks for the explanation
Taking hemispherical photographs in the field and post-processing the images	**AE**		All the reviewer comments are useful and should be considered as the proposal is revised. A critical aspect of using hemispherical photographs for modelling below-canopy radiation is the determination of the threshold for binarizing the images. In some of my research (e.g., Leach and Moore, 2010, Hydrol. Proc.), we have taken solar radiation measurements at photograph stations to allow selection of an optimal threshold based on matching observed and simulated radiation. I would recommend that the authors consider making some radiation measurements at some of their sites – even short-term measurements during field visits – as a check on the validity of the simulated radiation. If this approach is adopted, I would note that a thermopile-based pyranometer should be used under forest canopy, not a less-expensive photovoltaic sensor.	Thanks, related methods for photo post-processing have been added to the text.	
Taking hemispherical photographs in the field and post-processing the images	**AE**		In my experience, one cannot always guarantee suitable conditions for taking hemispherical photographs, as was emphasized by reviewer 1. For example, travel time to/from a remote site might not allow photographs to be taken pre-dawn or post-sunset, and one cannot guarantee uniformly overcast conditions. The following article describes software that incorporates an adaptive thresholding approach, which can be useful when lighting conditions vary substantially over the sky dome. Jonas, T., C. Webster, G. Mazzotti, J. Malle. 2020. HPEval: A canopy shortwave radiation transmission model using high-resolution hemispherical images. *Agricultural and Forest Meteorology* 284: 107903, https://doi.org/10.1016/j.agrformet.2020.107903.	Thanks for the paper. Agreed that hemispherical photo collection will require careful coordination. A series of photos with different exposure levels will be taken at each photo point to help alleviate this issue (related methods have been added to the text).	
Taking hemispherical photographs in the field and post-processing the images	**1**		Taking good hemispheric photographs in streams is slow and cumbersome and weather-dependent. Experience is very helpful. I'm concerned about the time and manpower it will take to get the photographs taken for this many sites and manipulations over such a broad geographic area. How many teams will be taking the photos? If more than one team, they need to be trained together and some comparisons done of cover and shade measured by the two groups on different days (when they aren't together). I would recommend that the authors conduct a feasibility study to see how quickly and reliably the photo surveys can be done.	Agreed. This study will take place over multiple years. Each year would likely focus on a subset of sites within a particular geographic region. That is, we are not planning on completing the entire state and all 20 sites in one year. Ideally there will be one highly trained photo team dedicated to working one site at a time. We anticipate developing these details in a forthcoming implementation plan and SOPs.	
Numbers and spacing of hemispherical photographs	**1**		The authors need to re-evaluate the spacing of the hemispherical photographs. Spacing should be determined by considering the scale of patchiness in forest floor light or by considering the scale of channel variation, typically one channel width. The current spacing of 7.5 feet seems to be arbitrary, and it is very short. It needs a rationale and justification. I worry that there will be little independence of measurements taken that close together. Garner et al. (2017) used 5m spacing, but they didn't justify it either. I suggest looking in the papers by Dana Warren's group from Oregon State because they have been looking at light patchiness in streams. I suggest a prototype effort to collect hemispheric photographs in these channels to better estimate the time and personnel requirements to conduct this work.	The 7.5 foot spacing of photos was determined based on the camera viewshed. The 7.5 foot spacing ensures that shade measurements (hemispherical camera viewshed) for a given plot will not be influenced by areas outside of the plot (see Figures 3a and 3b). Wider spacing would introduce influence from outside the treatment plot. Relevant language has been added to the text to explain photo point spacing. Following completion of this study design, we will be developing a detailed implementation plan. Part of developing that implementation plan will be to assess the time and personnel requirements needed to execute each step of this study.	
Numbers and spacing of hemispherical photographs	**AE**		I concur with the concerns expressed by both reviewers. Leach and Moore (2010, Hydrol. Proc) took 97 hemispherical photographs along a stream reach and found substantial intersite variability in simulated solar radiation.	The number and arrangement of photo points in this study was evaluated by an outside consulting statistician with expertise using hemiphotos to conduct similar studies. We acknowledge that slight changes in the position of the camera could have variable effects on shade estimates. Instead of relying on 1 centerpoint photo per treatment, we elected to take 5 photos per treatment, including plot center plus 4 additional points spaced at even increments from plot center. This will help ensure that we can capture the mean condition for each plot/treatment.	
Numbers and spacing of hemispherical photographs	**AE**		An additional consideration is that shading can vary systematically with position across the stream. I strongly recommend that the authors follow up on the reviewers' suggestion to conduct a pilot study to understand better the spatial variability of effective shade as a basis for developing a defensible strategy for the numbers and spacing of photographs.	See above response. We will use the mean of the 5 photo values.	
Numbers and spacing of hemispherical photographs	**1**	Pg. 9, Experimental unit section, last bullet	Why five photo points? How will the photo points be spaced? Why 5 feet from bankfull edge as opposed to channel center? Minimum channel width in site selection is only 5 feet. If the channel is relatively wide, say 25 feet (the maximum in the site selection criteria), will two or three photos need to be spaced across the channel? (comment from within study design doc).	Please see responses above and below. The number of five photo points was selected based on the recommendation of a consulting statistician with expertise with this type of research. The methods state: "The photo points will be located at a consistent distance from the plot boundary at a manageable depth (~<1 foot deep), to be determined after study sites are selected." The earlier statement, "Photo points located 5 feet from the bankfull edge of the stream/stream-adjacent plot boundary (see Figures 3a and 3b)" refers to the distance used to calculate shadow lengths - not actual field protocols (described later). The group discussed potentially locating photo points at channel center, but ultimately decided to use a consistant distance from bankfull edge rather than channel center, which will likely vary from site to site.	

Numbers and spacing of hemispherical photographs	**1**	Pg. 14, first sentence	What's the rationale for this distance (7.5 feet apart)? This is so close that the different photo values may not be independent. (comment from within study design doc).	The 7.5 foot spacing ensures that shade measurements (hemispherical camera viewshed) for a given plot will not be influenced by areas outside of the plot (see Figures 3a and 3b). Wider spacing would mean that trees, etc. originating outside the plot would potentially influence shade estimates for the specified solar period of analysis. We do not expect the five photo values to be independent. We recognize, however, that moving the photo point even short distances (e.g., 1-2 feet) could change the single photo values substantially, for example, due to one branch or canopy gap. Therefore, we will take multiple (overlapping) photos from slightly different perspectives 7.5 feet apart and use the mean of those values to calculate change in shade among the sequential harvest treatments.	
Numbers and spacing of hemispherical photographs	**1**	Pg. 14, first sentence	Several studies have shown the importance of streambank vegetation for shading channels. I suggest shooting the photos twice, once at 1m height, and once as low as possible. The amount of shading provided by bank vegetation is important. (comment from within study design doc.)	The group discussed this option and decided that one photo height at 1 m was sufficient for meeting the study objectives. We will analyze *change* in shade in response to the treatments, *all things being equal*, including streambank vegetation.	
Numbers and spacing of hemispherical photographs	**1**	Pg. 14, hemispherical photo collection sequence	This is a lot of hemispheric photos. Has anyone done some test runs to check feasibility and manpower needs? (comment from within study design doc.)	The number of hemiphotos was determined based on feedback from an outside consulting statistician (see related responses above). We have reached out to collaborators who are experienced in hemiphoto work to help determine time and personnel needs. The budget and timeline for hemiphoto work will be informed by these estimates.	
Forest and site characteristics	**1**	Pg. 12, site attributes section, first sentence	We need more data to test canopy cover and shade estimation models. There are several apparently good shade calculation models that incorporate tree height, overhang, canopy density, bank height, latitude, azimuth, and sun angle dynamics by date and time. Since a field study cannot cover the potential range of stream and canopy conditions, managers must rely on shade models to predict shade responses to treatments, so a good use of the generated data is to test and evaluate shade models.	Agreed. Although testing and evaluating shade models is outside the scope of this study, the data from this study will be available for those efforts in the future.	
Forest and site characteristics	**2**		Many of the geomorphological site attributes noted in this section would be good to record for general site characterization, but it's not clear how they would be used to help interpret pre- and post-treatment shade results, especially since shade is being defined based on relatively high solar elevation angles. It should be noted specifically how these variables will be used – e.g., in a formal statistical analysis or as general diagnostic information.	Updates have been made to the Analysis/Model selection sections to address this comment.	
Forest and site characteristics	**2**	Pg. 19, #2	This should include stand metrics for both pre- and post-harvest conditions.	No change needed. The study design will capture the pre-harvest stand metrics and shade conditions. #2 specifically refers to the *change* in shade *post*-harvest. This change in shade is relative to the pre-harvest condition, so it is necessarily described as post-harvest in #2.	
Forest and site characteristics	**2**		If possible, it would be good to estimate cover at both ~1 m and the water surface with a hemispherical densiometer for a more quantitative estimate of shade generated by low growing and understory vegetation and for consistency with other studies.	The group discussed this option and decided to only include one method (hemiphoto) and measurement height (1 m) in this study. See related comments throughout this matrix.	
Forest and site characteristics	**3**	Pg. 5, Stand Characteristics section, first sentence	Paragraph 1, first sentence: "Stand structural and compositional metrics known to influence stream shade will be measured in each plot for each harvest treatment and treatment zone (stream-adjacent no-harvest zone or adjacent-stand harvest zone)." For clarity and reproducibility, please explicitly identify the metrics "known to influence stream shade". If these metrics are the only tree metrics identified later in the paragraph, just say so. Please use language about factors, levels of factors and treatments carefully and accurately. Levels of factors are NOT treatments. Factors have levels and combinations of levels of factors comprise treatments. This error occurs in other places in the document and will need attention there too. The language in the first sentence of this section confuses these definitions. The sentence says 'measured in each plot for each treatment' which is confusing because (a) a plot has been defined as the large area which receives multiple no-harvest-zone-width levels (is a plot or subplot measured?) and (b) it does not explain WHEN the metrics will be obtained with respect to the timing of the implementation of the no-harvest-zone-widths. Something along the lines of 'measured in each plot after the implementation of each of the levels of zone width' would be clearer.	Text has been revised to remove confusing language.	Well done!
Forest and site characteristics	**3**	Pg. 5, stand characteristics section	It's stated that trees and tree measurements are 100% inventoried. Is this 100% inventory conducted once in each of the 3 plots on a stream or is it conducted once after each of the 4 implementations of the no-harvest-zone-width level within a plot? A suggestion for your consideration is to measure the tree density in the field after the various harvests. Variation in harvesting or even errors in the field can lead to tree densities that are not what you intended. Having a field measurement could be helpful if you find you have to 'adjust' for the actual density in the plot.	What Reviewer 3 describes matches what is being proposed. As stated in the document, "The tag number of each harvested tree at each treatment interval will be recorded so that stand characteristics (e.g., basal area by species) can be computed for the harvest and no-harvest zones for each interval." .... "After each thinning treatment, follow-up inspections will be conducted to ensure that all trees marked for harvest were harvested or at least felled. Additional harvest may be required to meet this standard. Additionally, any unintended tree falling or damage that occurred during the harvest activities will be recorded by tag number."	The revision on the new page 9 is more explicit - thanks.

Forest and site characteristics	**3**	Pg. 5, stand characteristics section	"Stand data will be used to help account for changes in shade in response to the treatments, variation in shade response among ecoregions, and the magnitude of statistical model variance. Stand data will be used to improve the fit of a Linear Mixed-effects Model (LMM; explained in the Data Analysis Procedures section) and control for site-specific conditions. Stand data will also be investigated independently of the LMM in relation to shade and treatment combinations." This description is written in a way that is not uniquely replicable. Is this a separate analysis from the LMM analysis that will "test" for differences among treatments? That is, is this a post-hoc analysis looking for potential explanations? Or is this a description of what will be done in the single LMM analysis that will formally test treatment effectiveness? If the latter, then the process for deciding which variables, and the function of each variable (linear, quadratic, log-scale) to be included (or not) should be described.	Under a newly-inserted subsection, Model Selection, describes the procedure for creating and selecting among models. This model selection procedure allows for comparison of different LMMs. We are not advocating for a post-hoc analysis that searches for potential explanations nor is this a description of what will be done within a single LMM.	Lines 607-608: "If they are important, they will assist with overall model fit....." I am unsure if this was meant as a statement about the unknown 'truth' (aka unknown statistical parameters) or a statement about the data analysis and estimated model. The statement is true if the authors meant the unknown truth but it's not true about the estimated model. In the analysis, adding any variable to the estimated model will always improve the fit - even if the variables are not important - so using model fit is not always the best way to choose a model. Lines 612-628 provides context but would be stronger with addtional details. The new section doesn't provide specifics about the decisions that will be made to identify the models that will be compared (Lines625 to 628) On what basis will models be proposed or the covariates to be included in the core model identified in the a priori process? This isn't included in the text. Since *a priori* hypothesized models/covariates are based on previous research so the models could potentially be identifed now and included in the proposal. If that's not possible some description of **how** they will be identified prior to examining the data would be best -right now it just says they will be identified. I refer the authors to a later Burham, Anderson and Huyvaert paper : Behav Ecol Sociobiol (2011) 65:23–35 that describes both *a priori* and *post hoc* use of AIC. If models will be proposed after looking at the data that process should also be described. With respect to proposing a prior models, the idea is to find a subset of models smaller than the set that includes all possible models or even most models. I appreciate the last sentence in this section - it's very informative, thanks.
Forest and site characteristics	**3**	Pg. 5, stand characteristics section	*"Stand data will be used to improve the fit of the LMM and control for site specific conditions"*. Explaining changes in shade with stand data is not the same thing as improving the model fit. Model 'fit' (observed data minus modelled response) will always be improved by the inclusion of any additional variable, so specifying the amount of allowable lack of fit is needed if you're trying to improve model fit. But I don't think you really intend to find a 'best fitting' model so much as want to know if stand data are correlated (to what extent?) with shade. This step is exploratory and it may be that graphical analysis is adequate compared to attempting to add stand variables to the LMM (which is complicated). By including such variables you may reduce the variation but induce bias and overfitting the bias/precision tradeoff in modelling. You may need to revise your wording here.	The text did appear to indicate that a post-hoc analysis will be performed, which is not our intent. The text has been changed to "Stand data will be used to control for site specific conditions." Also, please see the new text referenced in the above response. The model selection procedure avoids overfitting models and post-hoc data exploration.	See my comment on line 62. There are not enough details in the new model selection section to identify on what basis or which models will be chosen a priori. It just say they will be chosen. The revised sentence is better but it suggest that ALL stand data will be used. If there are specific site conditions which need to be controlled for they should be identified. 'Site'' is in your LMM as a random effect and will account for some of that variation. It would be helpful to identify what what conditions should be separately accounted for over and above that.
Forest and site characteristics	**3**	Pg. 5, stand characteristics section	Please be specific about how you will 'control for site specific conditions'. This is generic, there are many options for how this could be done, so this language is not specific enough for someone to repeat what you plan to do. Please also specify what you are controlling for. There will always be differences among the sites and plots that exist by chance so identifying what needs controlling for ***a prior*** i is important. Otherwise, it borders on data snooping.	Agreed. Please see above.	The text is more clear but I don't see explanation for which variables will control for which site specific conditions, or alternatively for how it will be determined if a condition has to be controlled for.
Forest and site characteristics	**3**	Pg. 5, stand characteristics section	The sentences refer to 'statistical model variance' (singular). In the LMM there will be many types of variances – variance among sites, among plots and among the smaller units within plots that receive zone width levels. I don't understand what is meant by accounting for model variance using stand data. Please explain what you intend to do.	Changing to: "Stand data will be used to help account for changes in shade in response to the treatments~~,~~ **and** variation in shade response among ecoregions~~., and the magnitude of statistical model variance~~	Excellent - thanks.
Forest and site characteristics	**3**	Pg. 7, general predictions and figure 3	Figure 3 and the text for the general predictions are not congruent. There are multiple versions of Figure 3 that would match the general predictions. I don't know which prediction (text or Fig 3) is the one you intend. Are the magnitudes of the differences among the zone widths for each harvest intensity also what you are hypothesizing (e.g. an exponential decrease in shade with over zone widths)? Please clarify.	Figure 3 and related text have been deleted.	Ok.
Forest and site characteristics	**3**	Pg. 7, Figure 3	Figure 3 is hypothetical, so I assume the error bars are also hypothetical. Are you trying to depict that you predict that the standard errors of the estimated treatment means will be constant? They may not be. If you estimate the mean for the 100' zone width it will have 3 times the replication of the other widths so its SE will be narrower. If this is hypothetical and really is a depiction of the TRUE means, no error bars are needed. Please clarify.	Figure 3 and related text have been deleted.	Ok
Forest and site characteristics	**AE**		Based on the reviewers' comments, the authors should strive for greater clarity in specifying what site characteristics will be measured, how/when they will be measured, and how they will be used in the analysis.	Revised accordingly.	
Forest and site characteristics	**AE**		Following from reviewer 1's comment, it would be useful for the authors to review a range of shade models to ensure that the relevant input variables would be collected. That would allow the data from this study to be used not only for the specific analyses of this study, but also to provide a test of shade models that could be applied in an operational context.	Testing shade models is outside the scope and budget of this study.	
Forest and site characteristics	**AE**		An additional comment is that bankful width and depth would be measured at the midpoint of the plot. The authors should provide more detail on the measurement protocols – e.g., how many cross-sections will be measured and their spacing.	Details about measuring bankfull depth and bankfull width have been added to the text.	
Forest and site characteristics	**1**	Pg. 12, site attributes section, first sentence	Add bank height. (comment from within study design doc)	No change needed. This seems similar to bankfull depth which is already included in the list.	

Study design - site selection and study population	1	Pg. 7, Study Population section, last sentence in first paragraph	Attempt to achieve a mix of widths? What if all the streams in an ecoregion are near the lower or upper size limit? (comment from within study design doc).	This study will be conducted on commercial forest land, so we are limited to sites where there are willing landowners and will be harvested during the study period. As long as the channel widths meet the site selection criteria, they will be candidates for inclusion in the study. Photo points will be located at a consistant distance from the streambank regardless of channel width. The channel widths included in this study will be recorded and described in the final report.	
Study design - site selection and study population	1	Pg. 8, Site selection criteria, second bullet	Any criteria for upland stands? Simply commercially valuable timber? (comment from within study design doc).	The site selection criteria specifically apply to the riparian stands.	
Study design - site selection and study population	2	Pg. 21, Site availability and sample size section	Identification of acceptable sites is likely to be a major challenge and potential limitation, and it may be worthwhile to consider sites within other ecoregions as noted on p. 21 in order to encompass the broad range of forest types within the state, rather than limiting the study to just 4 ecoregions. Suggestion: Note why the 4 ecoregions were specifically selected, or why the East Cascades and North Cascades (and maybe Blue Mtns) were specifically omitted. Is this due to relatively low levels of harvest in these regions? Given the likely challenge of finding acceptable sites to conduct the study, it may be worth considering sites in these areas to increase the number of options in the sample population.	The text has been modified to more clearly state: "These ecoregions represent the areas and environments where commercial timber harvest most commonly occurs in the state of Washington." The other ecoregions Reviewer 2 mentions have low levels of harvest and would likely be difficult to find sites that meet the criteria AND are being actively harvested during the study period (we did a GIS analysis of harvests in the state to determine the four selected ecoregions). The project risk section discusses the potential limitations to identifying acceptable sites for this study: "It may be difficult to identify an adequate number of sites that match the study population in areas where there are willing landowners or from approved Forest Practices Applications (FPAs) that will be harvested during the study period. Further, there is a small possibility that landowners may later choose not to harvest certain areas if timber markets are not favorable. To increase the number of potential sites, sites containing discontinuous plots (plots that do not share a boundary) could be considered for inclusion in the study, as long as the site layout does not introduce any unintentional biases that could affect outcomes. If five qualifying sites cannot be identified in one or more ecoregions, other options will be considered, such as: adding more sites in a subsequent year, continuing the study with fewer than five sites in an ecoregion, adding more sites to another ecoregion, removing an ecoregion from the study, substituting one of the four selected ecoregions with another relevant ecoregion in Washington, or adjusting the site selection criteria to include more sites. The study will include at least four sites per ecoregion and will only adjust site selection criteria if the criteria changes are carefully considered."	
Study design - site selection and study population	2		For the site selection criteria, it would also be good to note whether streams with low overstory density in the zones closest to the channel will be excluded, and/or what the specific density threshold would be for a site to be excluded.	There are no plans to exclude sites based on overstory density thresholds at this time. Given the difficulty of finding qualifying study sites, we do not want to limit our site selection possibilities or ability to reach sample size requirements at this step (also see other author responses throughout).	
Study design - site selection and study population	2		The use of GRTS to ensure more dispersed sites is interesting, but it's not as necessary as identifying sites with biophysical and topographic conditions that best meet the study objectives, so I'd suggest emphasizing that the site conditions will take precedence over spatial dispersion.	Agreed. See text for updates. Only sites meeting the selection criteria will be included.	
Study design - site selection and study population	2	Pg. 7, Figure 3	Note that the change in shade on the y-axis is a decrease in shade, since larger values represent larger reductions.	Correct. Figure 3 and related text have been deleted.	
Study design - site selection and study population	2		Define or describe what site classes II and III are, and why the study is being limited to these classes.	Added explanation. Site Classes II and III represent the conditions where timber harvest most commonly occurs. Table 1 has been added to further explain the site classes.	
Study design - site selection and study population	2		For stream widths, note that whether they are bankfull width or wetted width under low-flow conditions.	Text has been edited to indicate the stream widths are based on bankfull widths. This is also stated in the Site Selection section.	
Study design - site selection and study population	2	Pg. 8, site selection criteria	Somewhere in this section include some hemispherical images overlain with sun path diagrams to highlight the segments of the images that will be used for the effective shade estimations. This would help to communicate the study specifics to readers who may not be experienced in shade estimation using hemispherical photography.	Additional explanation has been added and the figure has been updated.	
Study design - site selection and study population	2		The primary window for field sampling and harvest is specified as between Jun 1 and Sep 1 which is a good general window, but it is recommended that researchers should have the discretion to expand or contract the sampling window, since warmer locations will likely be fully leafed out for a longer period, whereas colder locations may not fully leaf out until mid or late June. Having as much flexibility as possible will likely be needed given the challenging logistics of this study.	Agreed, revised accordingly. Note that this change only applies to when the sites are harvested. The hemiphoto analysis will still be applied within the 1 June - 1 September period.	
Study design - site selection and study population	3	Pg. 8, Site selection criteria	The listed site selection criteria on page 8 leave out a very important criteria - the landowner must agree to participate in the study. Please add this to your list. There are many reasons why a landowner may decline to participate – some of which could bias your sample away from the set of all streams in the targeted ecoregions that meet the criteria.	The site selection criteria is based on technical specifications. The text acknowledges that we must have landowner cooperation for access to study sites.	An important assumption, and likely an accurate assumption is that landowner cooperation is not correlated with shade conditions.
Study design - site selection and study population	3		Since you note that it's possible that there could only be 5 such sites in an ecoregion that match the selection criteria, the scope of inference could be limited – at least by the site selection. Conclusions that infer that results apply broadly would be incorrect and therefore statements about the presence or absence of effects should be carefully worded.	See below. Text has been added about the scope of inference.	Thanks.

Study design - site selection and study population	**3**	Pg. 15, QA/QC section, first bullet. Pg. 20, project risk analysis section, study scope paragraph	I suggest you make the scope of inference explicitly clear, that is, to what population the results of this study will apply. On page 15 it says that field inspection will confirm that sites are representative of the study population, but this population isn't clearly and completely identified in one location in the text. On page 20 it says "The findings may be interpolated within the range of the treatments but cannot be extrapolated outside of that range with great confidence" but this is not strictly correct. The findings will only apply to the environmental conditions **and** treatments in this study. Pages 20-21 do not refer to the limitations imposed by the site conditions and sampling of sites – it probably should.	See Study Area/Study Scope/Risk Analysis sections for related changes to site selection process and scope of inference.	Thanks for the addtion.
Study design - site selection and study population	**3**		Please clarify if there is a minimum distance between sites, whether or not different sites can be on the same stream, or within the same watershed. GRTS will help spatially disburse if you have many potential sites but what will you do if the geographic extent of potential sites is severely limited? Will you allow sites to be very close together? Some description of that situation would be helpful.	The text has been modified to clarify the randomized site selection process. There is no minimum distance between sites. The Project Risk section addresses your remaining questions.	Additions addded clarity - thanks.
Study design - site selection and study population	**3**		I appreciate the discussion of all the ways that sites can vary. It's important to acknowledge that as you constrain the sites to be more similar you are reducing the range of variation over which you can draw conclusions. This is a never-ending tension between enough variation to allow broad conclusions with associated large sample sizes versus narrow variation with narrower scope and smaller sample sizes. So, making sites similar will improve your precision but will decrease your scope of inference.	Agreed, thank you. Relevent text has been added.	Thanks. Having the figure include the intervals helps.
Study design - site selection and study population	**3**	Pg. 8-9, Experimental unit section	The plots are identified as the 'experimental units'. In this study plots are experimental units for the levels of thinning intensity and sections of plots are the experimental unit for the zone widths. In a strip or split plot there are multiple types of experimental units. The first 2 paragraphs of the section are about the plots, so my suggestion is to not refer to experimental units at all. See the figure at the end of this document that schematically identifies the different sized experimental units within this study. A defining feature of split and strip plot designs is the existence of different sized experimental units.	Agreed, revised and reorganized related text accordingly.	Thank you. The revision is much clearer
Study design - site selection and study population	**3**	Pg. 12, sample size section	The text does a good job explaining numbers of physical 'things' (regions, sites and plots) in the study and this is useful. The number of treatments needs revision because the document confuses the fact that the 100' width doesn't have an adjacent thinning intensity with the fact that all levels of width are assigned with plots assigned to a thinning intensity. So, there is a combination of the 100' width inside the plot with a thinning intensity of, for example 40. If the analysis will be done by with one LMM that includes ecoregion, then it may be helpful to accumulate replications across ecoregions. Note that the number of physical units isn't the same as the number of values (replications of an effect, aka a difference) used in the statistical analysis. So, the statistical replication (what is reflected in degrees of freedom) will be different.	The text has been changed to reflect that the main LMM analysis will not include the 100' no-harvest buffer data; those will be analyzed in an ancillary and simplified LMM. It is true that the number of physical units does not equal the number of values in the LMM. If the project proceeds with a strip-plot design then LMMs with the strip-plot analysis structure will consider the data set as having effectively three sample sizes based on the random effects structure.	Thanks, revisions are good.
Study design - site selection and study population	**3**	Pg. 12, last sentence in site layout section	Last sentence in Site layout section on page 12: *"The three different plot-level treatments (harvest sequences) will be randomly assigned to plots 1, 2, and 3 within a site, and will not necessarily be assigned in the order depicted in Figure 1.* " I found this sentence confusing (treatments confused with levels and not clear relationship between thinning intensity and zone widths. I suggest "One of each level of thinning intensity (40, 20, 0) is randomly assigned to each plot at each site. Within each plot, the specified level of thinning intensity is applied sequentially in time to create a sequence of the no-harvest zone levels. The sequence of no-harvest zone levels is in the same order in all plots." I also suggest adding this sentence to the caption in Figure 1.	Revised text accordingly to improve clarity.	Thanks, the revision is much clearer
Study design - site selection and study population	**3**	Pg. 13, sentence that begins "Bankfull width"	"*Bankfull width and bankfull depth will be measured at the midpoint of each plot".* Please rephrase this. The midpiont of a plot is a 2-dimensional measurement and not in the stream.	Revised text accordingly.	Thanks.
Study design - site selection and study population	**3**	Pg. 13, understory Veg cover section, second paragraph	"*A set of four oblique digital photos will be taken from the central photo point associated with each plot (Figure 1) to provide a visual record of site attributes, including understory vegetation cover (Table 3). Four photos will be taken from each point at 90° intervals (upstream, downstream, left bank, and right bank).* " Please include the temporal context for the photos as well – WHEN are the photos taken? This is an example of attending to spatial issues without needed temporal context.	Revised text accordingly.	Thanks.
Study design - site selection and study population	**3**	Pg. 13/14, effective shade section	"*The photo points will be located at a consistent distance from the plot boundary at a manageable depth (~<1 foot deep), to be determined after study sites are selected* " Please clarify what is meant by 'depth' or 'deep'. Depth is usually measured vertically but I think this is referring to a horizontal measure but that is unclear. Sentences further along describe the height above the stream at which photo points occur so I'm assuming this is not a measure in the vertical direction.	Thanks for catching that. Text has been revised to clarify this refers to vertical water depth.	Thanks.
Study design - site selection and study population	**3**	Pg. 14, stand characteristics section	"*After boundary marking……*" Would more clearly be "after plot boundary marking…… The information here should be coordinated with the information about the 100% inventory on page 5. See my other comments regarding this.	Revised text accordingly.	Thanks.
Study design - site selection and study population	**3**	Pg. 15, harvest layout and implementation section	A suggestion for your consideration is to measure the tree density in the field after the various harvests. Variation in harvesting or even errors in the field can lead to tree densities that are not what you intended. Having a field measurement could be helpful if you find you have to "adjust" for the actual density in the plot.	What Reviwer 3 describes matches what is being proposed. As stated in the document, "The tag number of each harvested tree at each treatment interval will be recorded so that stand characteristics (e.g., basal area by species) can be computed for the harvest and no-harvest zones for each interval." …. "After each thinning treatment, follow-up inspections will be conducted to ensure that all trees marked for harvest were harvested or at least felled. Additional harvest may be required to meet this standard. Additionally, any unintended tree falling or damage that occurred during the harvest activities will be recorded by tag number."	Thanks for pointing me to that.

Study design - site selection and study population	**AE**		As reviewer 3 points out, sampling environmental populations is challenging due to environmental variability that exists in a range of dimensions and inevitably involves tradeoffs. Based on both reviewers' comments, I recommend that the authors take another look at how they will frame the study population and how best to sample it to ensure that they achieve statistically significant results but also capture an operationally useful scope of inference. For example, as suggested by reviewer 2, developing a sampling frame based on forest type or biophysical site indices may be more effective than sampling within a fixed set of ecoregions, and might provide for a broader scope of inference.	See author responses throughout and text revisions. Ecoregions were selected primarily to ensure that study sites were distributed within geographic regions where commercial timber harvest most commonly occurs on the east and west sides of Washington state. There is no standardized or agreed upon stand typing system within our program, so the group decision was to use ecoregions. Given that it can be exceedingly difficult to find study sites for our research, we intentionally keep our options open at this point with fairly broad site selection criteria. We will be thoughtful about site selection after acquiring more information about potential/available study sites.	
Quality Assurance and Quality Control	**2**		It would be good to note how exposure settings for the hemispherical images will be determined. If the camera has a high dynamic range (HDR) setting it may be possible to automatically collect a series of images with different exposures to select the best image during the post-processing phase.	Agreed, thank you. Comment has been incorporated into the hemiphoto acquisition methods.	
Quality Assurance and Quality Control	**2**		It is stated that: "Data analysis procedures will follow standardized methods to ensure comparability with other studies and datasets." This statement is vague and should be more specific in terms of the specific standardized methods, studies, and datasets.	Agreed. Deleted this statement to remove vagueness.	
Quality Assurance and Quality Control	**3**		Please include direction on how to record missing data on the data sheets.	No change needed. As stated, "Field staff will be instructed to take detailed notes and photographs to document any anomalous situations." More detailed instructions are reserved for the forthcoming Implementation Plan/Field Manual/SOPs based on this Study Design.	Good to know that there will be formal SOP documentation proces and instructions that cover this.
Quality Assurance and Quality Control	**3**	Pg. 14, second paragraph	One bullet point says, "Hemispherical photographs will only be collected under uniform sky conditions when the sun is not directly in view, and according to the camera manual" and there is a somewhat different sentence on page 14 that says, "Photographs will be taken when no direct sunlight is visible, at pre-dawn, post-sunset, or under an evenly overcast sky." I suggest you decide which is most specific and use common language for both. There are conditions under which the sentence could contradict each other.	Deleted inconsistent/redundant sentence and kept the earlier, more specific sentence.	Thanks.
Quality Assurance and Quality Control	**3**	Pg. 16, eight bullet	"*Field datasheets (digital or hard copy) and digital photos will be duplicated and stored in a secure location as soon as possible following completion of each field survey*". This is a fairly generic statement which may suggest that this process is not well-defined. More details could generate greater confidence that is critically important step will be well-executed.	Deleted. This detail will be expanded upon in the forthcoming Implementation Plan/Field Manual.	Thanks. It wasn't clear to me that there would be a formal plan and process for documenting procedurs.
Project Risk Analysis	**3**	Pg. 21, study design assumptions	This section correctly identifies why the design is not a split plot design and the issues with assuming that it is. The strip plot analysis may help.	Reviewer 3's recommendation of a split plot design does not address the issue described in the section on study design assumptions. Both split-plot and strip-plot designs, conducted as intended, require randomized application of the two treatment levels. A strip-plot involves the application of one treatment, thinning density, within a plot. The second treatment, buffer width, must be applied by first starting with a 100' no-harvest zone and sequentially harvesting inward. We cannot stand trees back up and randomize the order in which we observe the effect of no-cut harvest width. This violation of randomization occurs regardless of whether we consider the study design a strip-plot or split-plot design and has important implications for interpreting study results. Reviewer 3 is correct that the study could also be analyzed as a strip-plot design. This situation of being able to choose between a split-plot or a strip-plot design is an artifact of our inability to randomize the buffer width treatment order. Since buffer width could not be randomized, it becomes a philosophical debate about whether the true form should be a split plot or strip plot design. Both have merit. For strip plots we sacrifice power for estimating both treatments (main effects) instead of one and improve our estimates of the interaction effects of the two treatments. We appreciate the recommendation and will adopt the strip-plot design, but the adoption of the strip-plot design does nothing to ameliorate the	Agreed. The split plot approach is sometimes chosen because analysts know how to do it and are familiar with it. It can be a reasonable approximation.
Project Risk Analysis	**3**	Pg. 21, Site availability and sample size section	Small sample size is a possibility. Please provide ***a prior*** i determined biological or management-relevant differences. Statistical significance is a function of the study design (larger sample size will always decrease the p-values) so it is possible that non-significant results are obtained even if the absolute magnitude of the change is relatively large or that a tiny change in shade could be statistically significant due to a large sample size. Therefore, it is advisable to ***a priori*** determine important changes and state that in the anticipated results. This way even a statistically non-significant value that is quite large can be identified and discussed. More importantly, if confidence intervals are wide and encompass both important and non-important values, it will be important to discuss the reasons for this and what is needed for future work.	The purpose of this study is not to determine whether change has occurred (based on an arbitrary p value level or a pre-defined management/biological difference) but instead to estimate the magnitude of observed change and associated uncertainty around estimates. This study represents neither effectiveness monitoring nor compliance monitoring. This study estimates the effects of riparian harvest treatments on stream shade. There are few similar studies and none with this level of replication. As such, there is uncertainty about how precise estimates will be and the site features that drive levels of precision. We are in agreement that we should discuss (and analyze, based on the data collection efforts described in the document) sources of observed error around estimates.	AIC will be used. Similarly to p-values, AIC can be affected by small sample sizes - that is, fail to detect differences among models when the models really are different. the proposed apriori model identification will help but identifying what makes a model be practically or biologically relevantly different could be helpful. No changes requested here but important to consider as you do the model selection.
Project Risk Analysis	**3**	Pg. 21, Site availability and sample size section	I appreciate the contingencies for adding additional sites. See my earlier comment about required spacing between sites. The same goes for plots. Please carefully consider how some of potential alternations might change your scope of inference and hence the usefulness of the study. Adjusting site criteria 'on the fly' can be particularly problematic if not carefully considered so I appreciate the comment that such changes will be carefully considered. But in the interest of reproducibility, further criteria should be identified ***a priori*** .	It is typically very difficult to locate study sites and landowner cooperators within our research program. A further limitation is that we are unable to select sites until after the study design is finalized and approved by the presiding committees. Thus, we intentially do not want to limit our options at this stage, to prevent falling short of our sample size requirements. However, we do plan to carefully select sites based on our criteria. Note that the 100% stand inventory data will allow us to describe the overstory conditions in detail.	Thanks for the clarification.
Project Risk Analysis	**3**	Pg. 21/22, variation in site conditions	I appreciate the discussion of all the ways that sites can vary. It's important to acknowledge that as you constrain the sites to be more similar you are reducing the range of variation over which you can draw conclusions. This is a never-ending tension between enough variation to allow broad conclusions with associated large sample sizes versus narrow variation with narrower scope and smaller sample sizes. So, making sites similar will improve your precision but will decrease your scope of inference.	Thank you, agreed.	Thank you.

Project Risk Analysis	**3**	Pg. 22, second paragraph	On p. 22 at the end of the 2nd full paragraph it says that the re-orientation of the streams to standard orientations (N/S, E/W) "*will ensure that shade response to the treatments is not influenced by differences in stream orientation across sites*". I've noted this earlier but it comes up here too - I question whether this is biologically reasonable. The under and overstory growth on a site is a function of the site and environmental conditions at a site and that includes the orientation of the stream relative to the tree stands. Re-orienting the image doesn't change the *in situ* conditions that affect plant growth and vigor. So, while some conditions may be standardize by the re-orientation, others are not. I suggest you clarify what conditions can and cannot be standardized.	The text has been re-worked throughout to clarify what non-treatment factors will be reduced/eliminated by photo post-processing procedures. The goal is to measure change in shade, all things being equal (by blocking by site), for our stream orientations of interest. The 100% stand inventory data will allow us to account for and explain differences in shade response that are related to overstory composition and structure, regardless of actual stream bank direction in the field.	Thanks.
Statistical analysis	**2**		Suggest that the authors consider using some sort of forest type metric that more effectively summarizes the biophysical conditions that control the sensitivity of shade change to harvest in the statistical modelling in place of ecoregions.	Ecoregions were selected because they provide a useful framework for distributing study sites across a range of geographic regions and biophysical environments in Washington. The 100% stand inventory data will allow us to describe and analyze the exact range of stand conditions ultimately included in this study. See text for updates.	
Statistical analysis	**2**	Pg. 18, third paragraph	On p. 18 it is noted that the linear mixed-effects model (LMM) will be used to develop equations for predicting shade based on canopy and site variables. Although the objectives and critical questions note that the relationships between stand characteristics will be examined, the development of predictive models is not explicitly stated as one of the core objectives. It seems that this should be noted as an objective, given the detailed canopy variables that will be collected and analytical approach that will be employed. As currently written, there appears to be a slight mismatch between the objectives and methods.	The development of predictive models is not an objective of this study. The LMM is somewhat different in the sense of model - the LMM is a statistical "model", or the analytical technique that will be used to estimate change in shade for this study.	
Statistical analysis	**2**		As noted in other comments, it is not clear why the untreated side of the streams will be masked from the analysis, and this is not necessary to assess the effect of hypothetical treatments on shade. A more appropriate approach that would be more reflective of actual conditions would be to simply estimate above and below-canopy global shortwave radiation and filter the data to exclude values for solar elevation angles below the specified threshold to reduce effects of neighboring treatments.	Excluding the untreated side of the stream from photo analysis reduces the amount of variability among sites that is not related to the treatments. Text has been added to clarify this point. The use of hemispherical photography methods for this study was a group decision.	
Statistical analysis	**2**		To reflect actual conditions more comprehensively, it would also be good to estimate effective shade for treatments occurring on the north banks of the study plots, even though the changes would not be expected to be as large as for treatments where direct shortwave radiation transfer would be more dominant.	This is addressed in multiple author responses throughout. Based on the context of this study, we chose to focus on the range of stream orientations where harvest treatments are likely to have the greatest effects on shade.	
Statistical analysis	**3**		The study design better fits a strip-plot analysis than a split-plot analysis. See schematic submitted on last page of reviewer 3 review comments.	See comments above.	Yes, thanks.
Statistical analysis	**3**		The references to factors, levels of factors and treatments are not always accurate	The document has been edited to provide better clarity. The word "treatment" unfortunately has different meanings when used to describe statistical and operational considerations.	Thanks for the clarifications. The new text is more consistent in its use of the words. I found it to be more clear.
Statistical analysis	**3**		Attention needs to be paid to the assumptions underlying the application of linear mixed-effects models, particularly diagnostic and remedial approaches (the latter in the case that the assumptions are not met).	Agreed. Text has been added: LMM assumptions will be tested following tests described in Pinheiro and Bates (2000). If assumptions are violated we will strive to correct them.	Thanks.
Statistical analysis	**3**	Pg. 18/19, Statistical analysis section	A general comment is that much of what is here has already be said earlier in the document but with different language and different emphasis and some details missing in previous sections. Earlier sections on analysis did not include all details and thus left the reader with gaps in understanding about the analysis and flipping back and forth between sections was needed. I suggest writing a single analysis section that incorporates all the parts using consistent language and terminology and then writing a summary of it if needed for the introduction. That would improve consistency in the language and help make sure that relevant details are provided earlier.	Agreed, the document has been substantially restructured accordingly.	Yes, revisions do a great job of clarifying
Statistical analysis	**3**	Pg. 18/19, Statistical analysis section	This section identifies the response as the CHANGE in effective shade. This is not consistent with earlier text or what is stated in the Objectives or Critical Questions on p. 1 which identifies effective shade as the response. Please reword to accurately describe the questions as understanding if the difference in effective shade responds to harvest intensity or zone width etc.	No change needed. The current wording is compatible with what is being proposed - 'change' and 'response' have very similar meanings in this context. For the purposes of readability, the purpose statement/objectives/critical questions use the more general 'response' and the exact definition of this ('change' in shade) is described as appropriate later in the document (along with a descriptive math equation).	I understand the authors' response. Since it can be confusing for readers to have to parse different meanings to 'shade response', 'shade','change in shade' and 'response' so my suggestion was intended to help the reader by removing the need for them to figure out precisely what was meant.
Statistical analysis	**3**	Pg. 18/19, Statistical analysis section	Given that the response is a difference from pre-harvest conditions, it is important to have measurements for the 100' width associated with EACH plot at each site as I noted in comments regarding Research Approach.	That is the plan. See difference in effective shade (ΔES) equation (individual plots denoted by 'j')	agreed.
Statistical analysis	**3**	Pg. 18/19, Statistical analysis section	Shade values will not be "normally distributed". Please explain why you are assuming this and why it matters. Proportions or percentages around 50% can be approximately normally distributed and the LMM doesn't require that the response be normally distributed, only that the various estimated residuals be normally distributed. Be sure you know how to check this assumption because the assumptions could fail even using the differences.	The text states, "Shade values will not be normally distributed; however, the differences in shade values will be approximately normally distributed." The text has been changed such that a mean, not median, of the five shade measurements will be taken for each plot. Means are normally distributed and differences in means are normally distributed.	Ok. But technically what is required is that the model residuals are normally distributed - the response is not required to be.
Statistical analysis	**3**	Pg. 18/19, Statistical analysis section	The LMM is described as having a fixed effect of ecoregion. This is reasonable. But it contradicts what is stated about analyses on p. 6 (Data will be summarized and analyzed according to ecoregion) and on p. 22 (Data will be analyzed according to ecoregion). Please correct or add to describe the multiple types of LMM's that will be used.	Deleted earlier sentence: "~~Data will be summarized and analyzed according to ecoregion.~~"	Thank you.
Statistical analysis	**3**	Pg. 18/19, Statistical analysis section	How will the constant variance assumption of the LMM be checked? The LMM analysis is going to a priori assume that the variation is constant over ecoregions or over sites. This needs to be investigated and the model should be altered to accommodate heteroscedasticity if that occurs.	The constant variance assumption will be checked by examining residual plots. If it appears that the assumption is violated, the models can be adjusted to account for heterscedasticity (see Pinheiro and Bates 2000). Relevant text has been added to the Analysis section.	Thank you.
Statistical analysis	**3**	Pg. 18/19, Statistical analysis section	For reproducibility of analysis, what is the analysis plan if assumptions for the linear mixed model fail?	If the assumptions are violated we will strive to alter the analysis so that they are addressed. This may include modeling heterscedasticity, splitting data, inclusion of variables, etc.	Thanks for thinking ahead.

Statistical analysis	**3**	Pg. 18, Analysis section, third paragraph	Paragraph 3 refers to the design as a split plot with blocking. That is a bit misleading because this entire design is more complicated than a text-book example of a blocked split plot design and as noted earlier it's not a split plot design. 1)Ecoregion is a fixed effect and the sites within ecoregion serve as the replication of the ecoregion (that is, 4 degrees of freedom to test differences among ecoregions). The ecoregion fixed effect exists in the LMM at a scale in the model higher up than the random effect referred to in the document as the 'block' and there is no blocking factor associated with the fixed effect of ecoregion. 2) The whole plot factor (thinning intensity) is randomly assigned to each plot within a site and the split/strip plot factor (zone width) is sequentially applied within each whole plot (plot). So yes, sites serve as a random effect (aka grouping factor, block, 'error' term) for the whole plot factor while the plots serve as the random effect (aka grouping factor, block, 'error' term) for the strip/split plot factor. 3) As mentioned earlier, the split plot factor (zone width) is 'stripped', not 'split' since the strip/split plot factor is applied in the same order in every plot. In a strip plot design there would be a random effect that accounts for the grouping of all plots at the time intervals (since the zones occur in the same order in all plots) which forms that basis for the test of the main effect of width, and there would also be a random effect (the final 'error' term in the model) on which the interaction of the zone and intensity is assessed.	We have changed the language of the document to indicate that models will include a strip-plot model (split plot would work too) parameterization with additional variables such as ecoregion included. Please see new text. The purpose of the second point is unclear, except to indicate that we are in agreement regarding the assignment of random effects (at least to a split-plot design). We believe the structure of the strip plot described in the third point is addressed in the new analysis text.	Thanks. My comments were only for clarity. Please see the sentence on line 524 to 535. It says that the model will include a random effect of the thinning treatment nested within the site (in addition to a random effect for plots nexted within sites). Thinning is a fixed effect - there doesn't need to be a random effect for thinning nested within site.
Statistical analysis	**3**	Pg. 18/19, Statistical analysis section	This is a complicated analysis. There is also the potential that some random effects may not be estimable if they are numerically confounded with other terms. So, I suggest involving a statistician to help program and interpret the analysis.	Thanks. A statistician has been consulted during the development of this study design, and has helped write and revise the analysis section. Continued assistance from a statistician will be considered for the analysis phase.	Great - glad to know.
Statistical analysis	**3**	Pg. 18/19, Statistical analysis section	This section suggests that ecoregion will be included in the LMM so that ecoregion and its interactions with zone and intensity are analyzed in one statistical model. However, on page 6 it says "*Data will be summarized and analyzed according to ecoregion*". This is different than what is in this section. Please determine which you will do and consistently explain that.	Deleted earlier sentence, "~~Data will be summarized and analyzed according to ecoregion~~."	Thanks.
Statistical analysis	**3**	Pg. 18, analysis section, last paragraph	The last paragraph on p. 18 says that you will 'test' if ecoregion improves fit. You can test if there is a statistically significant effect of ecoregion but that is not the same test as the test for model fit. I suggest you don't test model fit. The LMM is generated based on *a priori* identified design structure (including ecoregion) which means the LMM has inherent orthogonality build into it (that's why it's a 'design') and you should not remove any of the design variables from the LMM even if they are not statistically significant. This is not a setting in which model fit is an issue. Specifically, you say "Contrasts will be examined to statistically compare different treatments and treatments by area" - by which I think 'area' means ecoregion. It could be that the author used the word 'fit' when they actually meant statistical significance.	The sentence states: "[…] analysis will test whether including ecoregions in the model improves model fit by comparing models that do and do not include the ecoregion variable." Please see the response above and new text discussing model comparisons. It is absolutely feasible to compare the performance of several models which have different parameterization so long as they utilize the same dependent variable. The "test" is less a classical p-value driven test than an evaluation of model performance relative to other models. A model comparison utilizing AIC (or AICc, or QAIC, or BIC) takes into account degrees of freedom and model performance. Contrasts are a separate tool, available for estimating within models the difference between factor levels.	We agree. My points were (a) text says you'll test if adding ecoregion improves fit and that's a different test than testing if there's support in the model for ecoregion. I think you want the latter. Current statistical thinking is that variables that are part of the design should always be included in the model since they usually imply a grouping (correlation among observations in one group) or a restriction to the randomization.
Statistical analysis	**3**	Pg. 19, analysis section, #2 and #3	The analyses identified at the top of page 19 (#'s 2 and 3) make use of the site metrics (continuous variables). This analysis should also include the random effects used in the #1 analysis to accurately account for the various 'groupings' for sites within ecoregions, plots within sites and subplots (for zone width) within plots. Again, this could be complicated by numerical effects so involving an experienced statistician could be helpful.	This comment is interpreted as meaning that the same random effects structure as was used for analysis 1 is employed here, and not that the actual random effects are reused. The random effects structure for Analysis 2 may differ from Analysis 1. Analysis 2 seeks to examine how stand metrics relate to shade, not the different treatment levels. Therefore the random effects structure could be a random effect for site and another for plot nested within site. Analysis 3 is strictly exploratory and will make use of random effects as fit the questions being asked.	I think we're on the same page. My comment was interpreted correctly. If the response variable for analysis 2 is measured on the same 'unit' (plot with site) as the response in analysis 1 than the random effect structure would be similar. But if the response in analysis 2 is measured on a different type of unit (e.g. a subplot within a plot, or at the site level) the random effect structure would differ from analysis 1.
Statistical analysis	**3**	Pg. 18/19, Statistical analysis section	This section refers to 'predictive' equations suggesting that you intend to apply these results to streams and locations beyond the ones in your study. It is important to clarify if you intend to predict estimates that pertain to the streams you sampled, estimates for other very similar streams represented by your scope of inference (including owner permissions) or estimates for other streams not represented in your sample of 20 plots. The 'predictions' you generate, and more importantly the standard errors are a function of the estimated random effects in the model. Random effects are known to be poorly estimated unless the sample size is very large so concluding that the standard errors, confidence intervals or prediction intervals from the analysis apply beyond your sampled sites could be tenuous. You may want to consider not 'predicting' anything but simply reporting the estimates from your study.	Random effects within LMMs are assumed to be normally distributed and centered around zero. The analysis will verify these assumptions. Random effects estimates are specific to levels of nonindependence that the model accounts for and are not useful outside of the dataset at hand. We are in agreement on this point. However, the fixed effects estimates can be predictive. At the conclusion of the study the analysis may very well provide information on expected (predicted) responses for stands in certain regions and having specific characteristics. We strongly disagree with Reviewer 3 on this point. This study is part of the adaptive management program. The results of this study will indeed result in one or several potentially useful statistical models. These models form a basis of understanding system function. As the adaptive management program proceeds, these statistical models can be used to predict outcomes, and those predicted outcomes can be compared against future observations. Reviewer 3 suggests we consider "not 'predicting' anything but simply reporting estimates". We counter that by ignoring the information provided by this study CMER would squander an opportunity to learn from this study and refine future monitoring efforts.	The measure of precision for your estimate will depend on the type of prediction you are doing - different types will use different random effect variances in the estimation of the standard error. The random effect variance is poorly estimated with small sample sizes so the precision of some fixed effects could be small.
Statistical analysis	**3**	Pg. 18/19, Statistical analysis section	For reproducibility, please describe how the assumption that changes don't differ by initial shade will be tested, and also how the 'model will be adjusted accordingly' since there are multiple ways to do this.	We have changed the text to address these points.	Thanks - revisions are more clear.
Statistical analysis	**AE**		The authors should pay careful attention to reviewer 3's comments on the statistical analysis and should consult a statistician with expertise in strip-plot designs and their analysis via linear mixed-effects models.	We did consult with a statisician during the development of this study design. That statistician was the primary developer of the analysis section submitted for this review. They were consulted again to help address comments related to statistical analysis.	
Measurement of stream temperature	**AE**		Reviewer 1 suggested that stream temperature be measured. While I agree that it would be great to include stream temperature response to harvesting in this study, I am concerned that, depending on the timing of field visits and the execution of treatments, there may be inadequate pre-treatment data to allow the treatment effect to be isolated – e.g., if all of the field work is conducted in one season. From my own work, I have found that a year of pre-harvest data is the minimum for detecting effects of harvesting on stream temperature in a rigorous manner. I note that field work would be conducted over three years. The authors should consider the possibility of installing temperature loggers during year 1 at study reaches that would be treated in years 2 or 3, at least for a sample.	Stream temperature is outside the scope of this study. The sole focus on shade response was an intentional part of the study proposal, so that results would be produced relatively quickly compared to a multi-year temperature study. Given the limited duration (1-2 days) and scale (325 feet of stream) of each harvest treatment, and the sequential nature of the harvest treatments (multiple treatments occurring within the same plot over a short period of time), we cannot adequately address stream temperature effects within this study design.	

Measurement of stream temperature	**1**		This research proposes large-scale watershed manipulations over many sites. This is a rare opportunity to include other relevant measurements. Add-on studies are limited by the short duration of each manipulation step, but stream temperature could easily be incorporated into the study design. Temperature loggers are relatively inexpensive and easy to deploy. Temperature loggers could be placed in each manipulated reach and in a nearby reference stream several weeks prior to the first manipulation, and the pre-manipulation relationships could be used to characterize and quantify temperature responses to these manipulations. At the minimum, double cumulative temperature curves could be generated (akin to double mass curves), and non-parametric tests of the stepwise temperature ratios could be used to estimate statistical significance of responses. Other types of statistical analyses will be limited by the short duration of each manipulation, but graphical analysis could also be used to characterize how temperatures change in response to manipulations (e.g. are changes apparent to daily peaks? Daily medians? Daily lows?).	See above response. Also note that the individual plots are only 325 feet along the stream and occur adjacent to one another, so we could not rule out upstream effects from other treatments. Treatment levels will be randomly assigned to plots, so the ordering of treatments from upstream to downstream will not be consistent among sites.	
Measurement of stream temperature	**1**	Pg. 1, Purpose and objectives	Why is evaluating how stream temperatures respond not also part of the purpose? Measuring temperature responses is not an objective? (comment within study design doc)	See above responses. The study was not intended or designed to measure stream temperature. A key aspect of the study scoping proposal was acquiring shade results in a relatively short time frame for a relatively low cost. It is anticipated that future studies would build upon this RCS study, and potentially incorporate stream temperature.	
Re-measurement following treatment	**AE**		I have noted in a couple of thinning treatments of which I am aware that the retained trees tend to extend their foliage within a few years of treatment, and that understory vegetation develops, both of which have the effect of increasing shade relative to conditions immediately post-treatment. It would be useful to revisit the sites after, say, three years to re-measure shade. Therefore, it would be useful to ensure that the original photograph sites could be re-occupied. This could be accomplished by establishing survey reference points beside the stream and measuring the distance and bearing from the reference points to the photograph sites using a tape and compass.	Agreed that this would be interesting and the group has expressed interest in monitoring the longer term dynamics (e.g., windthrow and regrowth). Text has been added/reroganized regarding monumenting photo points so they can be revisted in the future. Note that only the final treatment level iteration could be monitored long-term.	
Literature review	**1**		I understand that a larger and more comprehensive literature review was previously developed in support of this study, but I will note a strong bias towards PNW stream temperature literature in this proposal. The physics of stream temperature and the geometry of canopy shade are the same everywhere on earth, so a broader review of the literature might reveal findings pertinent to this study design. The reviewer provided a short bibliography of relevant shade and stream temperature studies from outside the PNW, with a number of papers detailing temperature variation due to non-vegetative landscape controls. See RCS Review R1 doc submitted for actual list of references.	Thank you for the additional papers.	
Literature review	**1**	Pg. 2, Literature review	Reviewer states sentence in third paragraph "*Likewise, previous studies have detected measurable increases in stream temperature with reductions in shade of <10%*" is redundant with sentence in second paragraph "*Reductions in canopy shading of more than 6-10% have been associated with measureable increases in stream temperature*". See also a recent paper on understory shade removal: Raulerson, S., C.R. Jackson, Nathan D. Melear, Seth E. Younger, Maura Dudley, and Katherine J. Elliott. 2020. Do Sothern Appalachian Mountain summer stream temperatures respond to removal of understory rhododendron thickets? Hydrological Processes 34 (13): 3045-3060. (comment from within study design doc)	Thank you for catching that. Deleted second redundant sentence. Thank you for the citation. It has been added to the Literature Summary.	
Literature review	**2**	Pg. 2, Literature review	In this section where "density" and "tree density" are noted, it is advised to specify whether these densities are referring to stem density and/or the optical density of the crowns and/or canopy.	Clarified in the text that this refers to tree density.	
Literature review	**2**	Pg. 2, Literature review	In the review of stream shade, for completeness it would also be good to mention the potential importance of understory shade, dead shade, and topographic shade that all can be important at specific sites.	Comment incorporated into Literature Summary.	
General comment	**2**		The *Riparian Characteristics and Shade Response Experimental Research Study* seeks to fill a critical knowledge gap in understanding how stream shade loss is affected by different buffer width and thinning intensities in a tightly controlled study. This will potentially result in a fundamental and important scientific contribution to inform forest management activities and to help improve models of stream shade. The logistics of conducting the study are complex and very challenging and the Draft Study Design provides a well-written, clear, and very good initial draft of a research approach that seeks to balance the objectives with the considerable logistical challenges.	Thank you, agreed that implementation of this study will be logistically challenging. The author has communicated these concerns to the committee.	
General comment	**2**		It would be helpful if line numbers were provided to facilitate the provision of specific comments.	Agreed. Line numbers have been added.	
General comment	**3**		Thank you for the opportunity to review this study design. The investigation incorporates a complex study design with multiple spatial scales and temporal considerations, responses based on hemispheric photography, a proposed sophisticated statistical analysis (mixed model methodologies) and additional covariates. It is reasonably well-described but would benefit from revisions to more clearly explain the methodology with consistent terminology for the wide range of experts who are likely to read the description.	Thank you, revised accordingly.	The revisions are all excellent and the document is much more clear now, thanks.
General comment	**3**		This study is spatially and temporally complicated. There's a tendency to focus on either space or time in the various sections of the manuscript but not both in any one section. An overriding suggestion is to be sure that both space and time are incorporated in each section.	Agreed and thank you for catching that. The document has been reorganized to clarify the time and space components.	The revisions are all excellent and the document is much more clear now, thanks.
General comment	**3**		Please use accurate language around factors, levels of factors and treatments accurately - it is not always done in the document. A treatment is a combination of levels of multiple factors. I have more explanation the specific comments sections.	Revised throughout accordingly.	The revisions are all excellent and the document is much more clear now, thanks.

General comment	**3**		The response is identified as 'effective shade' but in the analysis section it is identified as a change in effective shade relative to pre-treatment values. This needs to be clarified and made consistent throughout the document. Most designed studies only estimate "changes", that is differences between treatments or between levels of factors. But if the response truly is a change from pre-treatment values then the language needs to reflect that differences between treatments are estimated after accounting for pre-treatment values.	Edits have been made throughout to improve clarity. The specific definitions are detailed in the Analysis section: "The main analysis response variable will be the difference between shade values for the nine different treatment level combinations (three no-harvest zone widths [the 100-foot no-harvest distance will be excluded] and all three thinning levels) and the original pre-harvest plot-level effective shade values. Shade values will not be normally distributed; however, the differences between mean shade values will be approximately normally distributed. The treatment level combination values will be subtracted from the original shade values to control for the initial differences in shade among sites." Difference in effective shade (ΔES) will be computed as: Δ⟦ES⟧_hijk=⟦ES⟧_hij0- ⟦ES⟧_hijk"	Thanks for the clarifications. The first of the 3 analyses (line 548) is to determine how treatments affect shade (not change in shade) and the description says that means will be estimated from the LMM. So should this section be revised to refer to change in shade rather than shade? Since the idea is that shade is not normally distributed but change in shade is, estimating shade from the LMM might not be appropriate.