*Response to reviewer comments on the manuscript “The Stream Quality Index: A Multi-Indicator Tool for Enhancing Environmental Management Communication” by M. W. Beck, R. D. Mazor, S. Theroux, and K. C. Schiff. Submitted to Environmental and Sustainability Indicators.*

*We thank both reviewers for providing detailed and thoughtful comments on our manuscript. Our responses to these comments are below.*

**Reviewer 1:**

Title

1. We don't know why "For Enhancing Environmental Management Communication" should be placed in the title? What is the Environmental Management Communication? Is the purpose of this article for environmental management communication?

*The intent of the SQI is to synthesize complex biological, chemical, and physical data for communication of results to decision-makers. We feel this addresses a critical need for environmental management communication and we describe the need for such tools in the introduction, e.g., lines 32-38, 50-54. Further, the prospectus for this journal focuses on promoting “research on indicators as drivers for environmental management”. Communication is an essential component of this need “guiding human actions to achieve sustainable management” (also from the prospectus). We feel justified in including this in the title.*

Highlights

1. About five highlights, we think these are general procedures and it is not the highlight of this research. We recommend to compare the difference with other articles, and evaluate the contribution, importance and creativity into the highlights.

*The highlights were modified to emphasize results and what the SQI provides.*

* *The Stream Quality Index (SQI) integrates biological, physical, and chemical data*
* *SQI categories summarized biological condition and likely vs. unlikely stressors*
* *Underlying data were preserved to help identify factors driving condition*
* *A web application was created to calculate SQI scores and access regional data*

Introduction

1. The introduction covers some references, however does not introduce the scientific problem. How will this study contribute beyond the current literature?

*This study contributes to the current literature by describing and applying an approach for synthesizing multiple lines of evidence for routine condition assessment. As we previously outlined in the introduction, this addresses a critical need within the management community because assessment tools that combine multiple indicators into a unified index are exceedingly rare (line 50). This is important to communicate results of complex indices to audiences that require this information for environmental decision-making. We feel the current introduction adequately describes the scientific problem.*

2. The motivation is not clear.

*Please see our response to the previous comment.*

3. Line 35-38, which shows that this article is an application purpose, not the purpose of scientific research.

*Yes, this article describes the technical foundation for the SQI that is intended for use in applied settings. We consider this both a scientific and applied exercise.*

2. Line 41-49, in addition to these indicators, it is recommended to quote some other important indicators.

*Additional citations were added: Rankin 1995, Karr 1999, Joy and Death 2002.*

3. Line 50-67, there are many integrated models or indicators. It is recommended to quote and explain the characteristics of other author's models and guide them to the scientific and academic motivation of this article.

*We believe that there are few integrated models or indicators that combine chemical, biological, and physical assessments. This was our primary motivation for this article. We welcome the reviewer to suggest any references that we may have missed.*

4. Line 66-67, is this the conclusion? This sentence suggests moving to the previous paragraph. What are the differences between the attempts and goals of this article, and the recommendations are guided.

*This sentence was removed. We elaborate in the discussion on how the SQI is an improvement over averaging (lines 391-403).*

2. Methods.

1. Line 74-76, why choose Southern California wadeable streams? It is recommended to increase the environmental background content.

*We use Southern California data to develop our proof of concept and emphasize that the framework can be applied in other regions. There is an immediate need in Southern California for an integrated index, as previously mentioned on lines 75-77 and 383-386. Considerations in applying the framework to other regions are described on lines 248-249, 464-466, and 499-501. Expanding the index outside of Southern California is a future application that will be pursued, but it is irrelevant for describing the index in the current study.*

2. Line 79, why only choose benthic macroinvertebrates, algae? others? fish and aquatic insects? other indicators?

*California has developed bioassessment methods for macroinvertebrates and algae and their inclusion in the SQI was appropriate. However, we suggest that alternative bioassessment response indicators could be included where they are available (e.g., alternative bioassessment endpoints, line 502). Our framework is flexible by design to include different indicators or stressor data that vary between regions or monitoring programs, as noted in our response to the previous comment.*

3. Line 81, why are there no other factors such as dissolved oxygen?

*We specifically chose TN, TP, and conductivity because they are widely measured in many monitoring programs and they are commonly associated with development gradients in the study region. These variables also act as surrogates for unmeasured water quality pollutants or their effects, such as low DO. Please see lines 149-158.*

4. The four narrative assessment categories? How to align with management processes? Management processes?

*Line 89 was revised for clarity: “The four narrative assessment categories were defined in a way that would align with management processes by describing biological condition and suggesting which stressors are associated with the condition. These categories provide a first indication of how biology at a site responds to stressors, which could then be used to prioritize follow-up actions, such as causal assessment.”*

5. Line 117-123, this paragraph is important. But we don't know these definitions and criterions, such as the 1st, 10th, and 30th percentiles of scores?

*The sentence was revised for clarity and citations to Stoddard et al. 2006 and Ode et al. 2016 were added: “For both the CSCI and ASCI, the 1st, 10th, and 30th percentiles of scores at reference sites with minimal human disturbance [4,5] were used to categorize all sites as very likely to have altered biological condition…”*

6. Line 125-136, this paragraph is crucial. The paragraph tell us about the integrated process of

the two organisms or indicators. Is there an objective approach? Also, if there are three species or more?

*This approach was guided by the independent applicability of each index (line 129), so that information from both indices were considered in the assessment of overall biological condition. The condition assessments provided by both indices were also based on percentile distributions of reference sites, as noted above. Given these two criteria, we feel the combination of the two indices is fairly objective. The involvement of the technical advisory group also ensured that the combined categories could be communicated to managers. We see no reason why a similar approach could be applied to more than two indicators.*

7. Line 144-148, the sentence is like the introduction and discussion.

*These sentences were removed.*

8. Line 157-158, the sentence is like the introduction and discussion. How to deal with similar problems or factors?

*This sentence is critical for describing our choice of water chemistry data for the SQI and also describes an important caveat. We have retained the content. Also please see our response to comments on lines 79 and 81.*

9. Line 165-170, About IPI and CRAM, our question is how to (five sands, fines, or concrete, Shannon diversity of aquatic habitat types, Shannon diversity of natural Substrate types, evenness of flow habitat types, and riparian vegetation cover).

*These metrics are well-described in the references provided (Rehn et al. 2018 in particular).*

10. Line 210-212, this is the highlight of this article. Is there a statistical hypothesis?

*This equation describes how we defined the likelihood of observing altered biological condition from both physical and habitat stressors. It is a deterministic equation and not a statistical hypothesis test.*

11. Line 225-228, our question is the four possible combinations of biology (4 categories) and stressors (4 categories)? Should be 4 x 4 = 16? How to reduce it to 4?

*These are summarized categories for the overall SQI that combine information from the separate biological and stressor categories. We have added a sentence for clarity: “An altered/unaltered condition could result from one or both biological indices and a stressed/unstressed condition could result from one or both stressor types.”*

Results

1. The paragraph of results seems to be a statistical text, it is inadequate and should be strengthened. In particular, it is recommended to write the contents of ecological phenomenon and meaning.

*Our study presents a tool for communicating multiple indicators in a synthesized format to environmental managers. This information is empirically combined following well-described stressor-response relationships for aquatic biota, with appropriate citations describing the theory and understanding in the methods. Our results present the application of the SQI to southern California streams and the discussion describes how the SQI can be used in practice, including limitations and comparisons with existing tools. As such, a detailed discussion of the ecological phenomenon for our specific results is not within the scope of this paper.*

Overall

1. The Stream Quality Index: A Multi-Indicator Tool for Enhancing Environmental Management Communication could be one of the original and rigorous studies of international importance.

**Reviewer 2:**

The authors present a predictive index for stream quality (SQI) that looks to integrate biological, physical, and chemical status, but the main innovation is the attempt to relate biology to stressors, including as well a web application to calculate SQI scores from regional data with the purpose to involve managers and wider audiences. This index that compares observed benthic macroinvertebrate taxa and metrics at a site to those expected under near pristine conditions are based on models (logit link functions) that estimate the likely macroinvertebrate community from environmental parameters/indices.

I consider the present manuscript a valuable and original work where, nevertheless, I call the authors attention for some difficulties in different steps of the process, which should deserve, at least some observations in the discussion.

The "four narrative assessment categories" of SQ are, in my opinion not very clear. E.g. "healthy and unstressed" or "healthy and resilient" are not easy to differentiate. The same difficulties arise between "impacted and stressed" and "impacted by unknown stress", especially when the authors insist with the potential use by different audiences. The separate categorical outputs created for the biological condition and stressor condition are also very confusing: this is, the ones resulting from high/low CSCI and high/low ASCI.

I also find the categories too narrow based on the percentiles of scores at reference sites. Is it possible to distinguish increasing degradation with SQI? This is, I doubt that the numeric values are in fact sensitive to assess differences between sites in marginal or extreme conditions… An obscure aspect for me is also the need that the technical advisory committee "with representatives from local management institutions" may provide guidance on assigning these values. Don't they introduce too much noise in the process if they are not qualified? Does it not make very subjective the stress thresholds?. Another aspect that is not treated is how far this index may be influenced by the typological conditions? It has only a regional character or can be extended to other geographical regions?~ Anyway I think SQI is more important for managers to determine how different stressors influence relatively the biological condition than this idea of communication to high-level or to non-technical staff… Considering the structure, I think that chapter 2.3 has to be shortened since indices IPI and CRAM were previously published....

Fig.1 could be provided as a graphical abstract too…

Other aspects that need to be clarified/ discussed:

Line 53. It is considered that a single unified assessment is preferable when communicating stream health to non-technical managers, but I'm not sure about this combination: it integrates sets of stressors but probably that process makes the index more complex and less sensitive to specific stressors…

Line 89. It is not possible to understand how the "four narrative assessment categories" were defined in a way that would align with management processes…

Line 105. The ASCI was developed as a response end point only for lower trophic levels: this aspect does not represents a constrain in higher disturbed levels?

Line 142. I don't think that habitat is a responsive metric, but on contrary, it can be described by metrics...

Line 308. I don't find it's interesting to distinguish patterns between ASCI and CSCI. In fact they should be analyzed together for SQI…

Line 402. There is a crucial question arising from this work: Why not then to integrate water and habitat quality in a more general ecological quality assessment?

Line 408. The idea of categories implies that they are defined for a certain region and cannot be generalized or applied in other areas… This aspect is not present in the limitations of the SQI approach..