*Response to reviewer comments on the manuscript “The Stream Quality Index: A Multi-Indicator Tool for Enhancing Environmental Management” 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 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”. We have removed “communication” from the title to more closely align with the goals of the journal.*

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*
* *The SQI was created for southern California but can be calibrated for other regions*
* *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. 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). We have added this sentence to the final paragraph of the introduction: “Our approach will contribute to the current literature by presenting a proof of concept for synthesizing multiple condition assessments in an easily communicated format, while retaining the individual components for follow-up analysis.”*

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.

*The addition to the introduction to clarify how this work contributes to the current literature provides a context for this statement. The sentence was also revised: “This is a likely improvement over treating indicators as equivalent and simply "averaging" results to assess overall condition.”*

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.*

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 human land development gradients in other regions. These variables also act as surrogates for unmeasured water quality pollutants or their effects, such as low DO, which are difficult or expense to quantify at large spatial scales. Please see lines 149-158. As mentioned in response #2, the SQI can be modified to incorporate the relevant stressors in other regions.*

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?

*We’re not entirely sure what was meant by “objective approach”. If the comment relates to objectivity of the methods for how we integrated indicators, then are approach was guided by the independent applicability of each index (line 129). This allowed information from both indices to be objectively 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, which are based on empirical results and implicitly objective. Given these two criteria, we feel the combination of the two indices is fairly objective. A similar approach could be applied to more than two indicators.*

*If the comment relates to the inclusion of the stakeholder group as potentially affecting our objectivity, please see are response to reviewer two (response to comment starting “An obscure aspect for me is…”).*

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).

*We have added the following sentence for clarification: “Detailed descriptions of each metric, including how they are measured, are included in Rehn et al. (2018).”*

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: “A healthy/impacted condition could result from one or both biological indices and a stressed/unstressed condition could result from one or both stressor types.” Also, please see our response to the first comment from reviewer 2 where we have elaborated on the SQI condition categories.*

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. Although a detailed discussion of the ecological phenomenon for our specific results is not within the scope of this paper, we have included some text in the discussion to clarify some of the relationships reflected by the SQI:*

*“To properly capture relationships among indicators of stream quality in a way that is consistent with conceptual modeling of a healthy stream ecosystem, it was crucial to develop an index that accurately reflected biology’s role as a direct measure of condition, and that reflects physical and chemical indicators as measures of stress. These relationships were captured using the stressor-response model that reflected an empirical and ecological relationship between biology and stressors. Empirically, we quantified the likelihood of different sites having altered biology across stressor gradients. Ecologically, stream biota have documented and mechanistic responses to the in-stream or riparian stressors included in the SQI [2, 19, 25, 26]. These linkages were reflected in the separate categorical components of the SQI; altered biology was more often associated with high stressors. Detailed outputs provided additional information about ecological relationships. For example, some sites were altered for the ASCI and had high water chemistry stressors, whereas other sites were altered for the CSCI and had high physical habitat stressors. This reflects a potential difference in the sensitivity of algae to water chemistry and macroinvertebrates to physical habitat, highlighting the utility of using different assemblages to assess condition.”*

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.

*As reviewer one also points out, the explanation of the categorical outputs was somewhat confusing. We have added clarity to better explain the assessment categories (starting line 228): “A healthy/impacted condition could result from one or both biological indices and a stressed/unstressed condition could result from one or both stressor types. The first SQI category is typically assigned to pristine sites (healthy biology, low stress), whereas the fourth category is typically assigned to degraded sites (impacted biology, high stress). The second and third categories are assigned when biology and stressors provide different information. A healthy and resilient site has unaltered biology, but stressors are present (i.e., the biota are resilient to stress). A site that is impacted by an unknown stress has altered biology, but neither water chemistry nor physical habitat stress is observed (i.e., biology is likely impacted by other stressors not included in the SQI).”*

*Additional content was added to the following paragraph describing the separate biological and stressor categories:*

*Line 232: “An overall healthy condition for the SQI only occurs if both the CSCI and ASCI indicate healthy conditions, whereas an impacted biological condition is assigned to the overall SQI if one or both biological indices is/are impacted.”*

*Line 238: “An overall unstressed condition for the SQI occurs if both chemistry and physical habitat stress are low, whereas a stressed condition is assigned to the overall SQI if one or both of the stressor types is/are high or the additive effects of both exceed the threshold.”*

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…

*We certainly agree that the categorical outcomes reduce sensitivity of the index to distinguish sites along condition gradients. Our goal was to create a communication tool that synthesizes multiple indices in a unified format (i.e., lines 53, 58). As the reviewer correctly notes, synthesis comes at the cost of specificity. Thus, the SQI was developed as a high-level screening tool to quickly compare conditions between sites and to develop priorities accordingly. A rapid comparison of multiple biological indicators and stressor data is not possible using standalone indicators and the SQI was meant to address this need. We do encourage readers to consult the component indices of the SQI that are available on the website if more specificity is needed.*

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?.

*Inclusion of the technical advisory committee was critical in the development of the SQI. The individuals were selected based on their expertise and were also intended users of the finished product. We preferred not to develop the SQI in isolation of the management community, but rather to include them as contributors to develop a shared sense of ownership of the product. This promotes the likelihood that our intended audiences will actually use the SQI to make decisions about management priorities. Further, some decisions in the development process were not strictly scientific decisions (i.e., specific categorical thresholds for stressed/unstressed) and required feedback from the advisory committee. This provided an avenue to incorporate the scientific/technical components of the SQI (e.g., stressor-response models) with the needs of the management community (e.g., where on the spectrum do we define an impact that has meaning for managers). Numerous studies have documented the inclusion of stakeholders as being critical in the application of science in decision-making (e.g., Pohjola and Tuomisto 2011, Env. Health 10:1-13).*

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?~

*Please see our response to reviewer 1 regarding application to locations outside of southern California (comments on lines 74-76 and 79).*

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…

*The SQI addresses a specific need within the management community to synthesis multiple indicators into condition categories. This allows the results to be readily communicated, which will likely facilitate management action, as opposed to inaction or misguided responses if results are poorly communicated. The web app also allows for a detailed examination of differential and cumulative stressors that can easily be identified for follow-up action.*

Considering the structure, I think that chapter 2.3 has to be shortened since indices IPI and CRAM were previously published....

*Some details about IPI and CRAM were removed: lines 174 and 182.*

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

*We have included a graphical abstract that includes a snapshot from the web application.*

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…

*Please see our response above to your general comment.*

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…

*This sentence was revised in response to similar comments from reviewer 1: “The four narrative assessment categories were defined in a way that would align with management processes by indicating biological condition and suggesting which stressor categories 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.”*

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?

*Yes, the ASCI describes algal condition with implications for nutrient-processing and effects on upper trophic levels. This was edited: “The Algal Stream Condition Index (ASCI, Theroux et al. in prep) was similarly developed as a biological response endpoint for primary production, with implications for ecosystem function.”*

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

*Revised: “Depending on the context, stream health or condition can be reflected by physical habitat as a response to external drivers”*

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…

*The ASCI and CSCI are response endpoints that describe different components of the biological community. Because biota respond differently to stressors depending on their ecology and environmental needs, the ASCI and CSCI could be diagnostic of different drivers of stream health. Our technical advisory committee also supported the use of both indices in the SQI and describing differential outputs could be of use to managers.*

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?

*As noted above, our intent was not to create a general assessment tool that simply averages multiple indicators. There is potential value in simplifying information for communication, while respecting the independent applicability of each component index. There are many cases where biology does not reflect stressors and vice versa and our combined index sufficiently describes these exceptions.*

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..

*Please see our responses above regarding application to other regions. Specific suggestions for applying the SQI beyond southern California are provided at several points in the text.*