*Response to reviewer comments on the manuscript “The importance of open science for biological assessment”, by M. W. Beck, C. O’Hara, J. Stewart Lowndes, R. D. Mazor, S. T. Theroux, D. J. Gillett, B. Lane, and G. Gearheart.*

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

**Associate editor comments:**   
  
I believe the manuscript addresses an extremely important topic that should be relevant to virtually all Freshwater Science readers. Overall, the arguments are convincing and the writing is adequate. Both reviewers and I have provided numerous suggestions that should greatly improve the manuscript. Reviewer 2 is a state-level bioassessment expert, and provides several additional questions/issues that, if addressed, could make your paper more appealing to that audience. Consider their suggestions—although some are a bit off topic. Reviewer 3 is from academia and provides many excellent suggestions to improve the flow and simplify your message. They suggest a change in the tone of your manuscript—to one that is less pontificating. I didn’t get the same impression—but largely because open science has been drilled into my head as a federal government scientist. I provide below a few general suggestions, followed by detailed notes as I read through the manuscript.  
  
Terminology is inconsistent. For example, “methods,” “tools,” “products,” etc. are used interchangeably and it leads to confusion. There were sentences where “tool” made more sense but you used “method” instead. If you want to keep using all these terms, you need to define them up front. Otherwise, use a consistent term that encompasses all these aspects of bioassessment.

*We have revised the text for consistency in terminology. In short, the text now refers only to “bioassessment products” and “open science tools”, with definitions for both provided in the introduction. Please see our specific response to your comment on Line 80.*   
  
Reviewer 2 was troubled by redundancy throughout the manuscript. On the one hand, redundancy is helpful when you’re trying to, in essence, teach readers a new thing. But you also run the risk of an excessively long manuscript, and annoying readers who feel like they’re reading the same things over and over. I’ve indicated in my specific comments where you might trim some material so that redundancy is less glaring (as did Reviewer 3). But I also urge you to take a close look at how you could minimize redundancy throughout the manuscript. Another idea to consider is that you seem to discuss the concepts of, say, open data in the “Principles” section, but then repeat a lot of the same material in the “Bioassessment” section. I suggest carefully wording each section so there is less redundant information.

*Our repetitive approach was somewhat intentional in order to emphasize key concepts, as you suggested. However, the manuscript is quite long, and we agree it can be shortened in some instances while still communicating critical ideas. We have removed portions of the text where you and reviewer three have commented specifically. As one example, the content from the section “Open science in practice” that described the SCAPE tool was shortened from ten paragraphs to four. See our responses below for specifics.*

*In total, these omissions have shortened the manuscript by*   
  
I’ve also indicated sections in the document that stray from the main topic (e.g., discussion of thresholds), and that are burdened by excessive detail (SCAPE tool description). These sections should be removed and / or dramatically trimmed.   
  
AE SPECIFIC COMMENTS  
  
Line 50: Not sure how “methods” could be transformed into actionable information. I can envision results and tools being transformed.

*Agreed, sentence was revised to “…that transform bioassessment results into actionable information.”*  
  
Lines 59-62: This sentence is too complex and also unclear. Consider splitting and rewording to clarify. Do you mean that bioassessment tools are being used in a post hoc fashion rather than for protective/preventative purposes (e.g., anti-degradation)? OK, the sentences after this one provide clarity & examples. DELETE THIS sentence.

*Sentence was deleted.*  
  
Line 67: Unclear what you mean by “discoverability of existing products by other research teams…” I sort of know what you’re getting at, but the sentence needs to be more clear. I think you’re saying that more discoverability and open sharing of bioassessment data, methods, and tools among researchers is essential to…

*Revised sentence: “To these ends, more open sharing and discoverability (i.e., improved database connectivity and ease of access) of bioassessment data, methods, and tools is essential to…”*  
  
Line 72: Have bioassessment tools really become “diagnostic?” To me, a diagnostic tool reveals WHY someone is sick. I don’t think we’re there yet with biological assessments. We do a good job of determining WHETHER a stream is impaired, but can you cite evidence that we’ve been able to develop tools that diagnose the potential causes?

*Agreed, most indices do not provide this information and causal assessments could be used to identify specific stressors. I replaced “diagnostic” with “sensitive”.*

Line 80: You seem to be using bioassessment “products,” “methods”, and “tools” interchangeably. I think you need to define what these mean, or stick with a single general term.

*We agree our use of different terms was inconsistent and confusing. We have chosen to use “products” for bioassessment and “tools” for open science. We feel this adds consistency of meaning and have included a brief description in the introduction to clarify each term. From the second paragraph: “In the United States, the CWA gives power to states, tribes, and territories for bioassessment development, where final products (i.e., assessment indices or other products that support biointegrity decisions) require federal approval…”. From the final paragraph of the introduction: “Herein, open science "tools" describe best practices and specific applications that use an open philosophy to support applied science.”*  
  
Line 87: Not sure what you mean by “…existing methods may not be discoverable beyond immediate research applications…” Are you referring to actual bioassessment tools such as an IBI for a region? Or are you referring to the actual methodology, which is typically “discovered” by reading the literature?

*Sentence was revised to: “Moreover, existing indices may not be easily replicated beyond initial research applications…”*  
  
Lines 90-92: Here you focus on “methods,” but I don’t see how managers would be interested in choosing methods. It seems like that is a challenge for the scientists. Again, are you referring to tools or applications?

*This sentence was revised in accordance to the general comments above about consistency of terminology: “The abundance of available products can be a point of frustration for managers given a lack of guidance for choosing among alternatives, particularly as to how different assessment products relate to specific management, monitoring, or policy objectives (Dale and Beyeler 2001; Stein et al. 2009).”*

Line 105: Publicly funded science is not unique to bioassessment, so I don’t see why this sentence is relevant.

*This was revised as we our original intent was to highlight the link to public natural resources, not public-funding: “Legal and ethical precedents in bioassessment may also necessitate open data sharing given that environmental monitoring programs are often established to protect and maintain publicly-owned natural resources.”*  
  
Lines 124-153: Nicely written!  
  
Line 174: Most synthesis products are, in my opinion, GREATER than the sum of the individual datasets. The SUM, would be just a summary of the individual datasets. A synthesis implies taking the additional step of advancing the science.

*Sentence was revised: “Collaborative publications have increased in the environmental sciences as research teams leverage open data to create synthesis products that allow novel insights from comparisons across multiple datasets”*  
  
Lines 202-258: This section should be reconsidered. In it, you provide an introduction to the sections that follow, but in so doing you discuss many of the ideas that will follow. So, when readers actually read the next section, they get the impression they’ve heard this before. I suggest dramatically reducing this section and dropping ideas and examples that you’re going to flesh out in the subsequent section.

*This section was shortened to retain only key ideas that support the introduction of the following sections. Lines 208-231 were removed – these lines described molecular advances in bioassessment, which are important but is unnecessary information for our larger goals. Lines 245-258 have also been removed as they are almost entirely redundant with the following sections.*

*We have retained most of the content on lines 203-214 and 232-244 because it provides a good lead-in to the following sections by 1) describing the overall workflows in figure 2 (lines 203-214), and 2) clarifies that openness occurs on a continuum (lines 232-244). We feel that this material is not redundant and emphasizes key points to provide additional context in the following sections.*   
  
Line 286: This is an example of a redundant idea. You talked about this in Line 250.

*Line 250 was removed.*   
  
Lines 352: Again, I get the feeling I’ve read this in previous paragraphs.

*This line was removed.*  
  
Line 364: I’m not sure what you mean by “ bioassessment translation” You describe it in the opening paragraph, but it still doesn’t seem like the correct word choice. Why not “application” or “implementation?”

*Agreed, “translation” is a bit loose and we’ve changed it to “bioassessment application” in the subsection title and other instances in the text. We also provide a clear definition of what this means in the opening paragraph: “Bioassessment application is the linkage of products between the research and management community, where links are enabled through open science tools that can deliver the products using a reproducible and accessible platform.”*   
  
Line 390: You’re using jargon (geometries, aesthetics) from Wickham here. The readers that you’re trying to convince to use these types of tools will not understand these terms.  
*Replaced “geometries and aesthetics” with “instructions”.*

Lines 526-537: This paragraph seems to be straying off topic. Here, you’re delving into the challenges of setting thresholds. Suggest dropping paragraph.

*Paragraph was removed*.  
  
Lines 538-552: Same as last paragraph. You’re getting into one technical challenge of bioassessments that is a good topic for another paper, but seems tangential to this one. Drop paragraph.

*Paragraph was removed.*  
  
Lines 570-639: Pointing to this tool is a good idea, but you go into WAY too much detail about its development. Reduce these paragraphs by 90% by simply stating the problem that the tool was developed solve, that it was developed using open science principles (no need for detail here), and how it has been received. Cite your forthcoming paper for the details.  
*Lines 553 – 639 were shortened substantially to highlight the key points of SCAPE that are relevant for understanding open science applications in a practical example. This reduced the content down to two paragraphs, one describing the problem that SCAPE addresses and another that describes the open science components. Many details for the latter were minimized to reduce redundancy with previous sections.*

**Reviewer 2 comments:**

This paper is ambitious in scope. It includes a philosophy, a catalog of tools, and a realized example of the philosophy and tools in action. The open science paradigm is a bold proposal that is certainly worthy of debate in the scientific community.   
  
The manuscript is well written, well organized, includes sound science, and communicates concepts clearly. However, it has some gaps, particularly in the area of the underlying assumptions of the open science philosophy and its possible real world implications. Because the novelty of this paper is philosophical rather than technical per se, and the authors are calling for a philosophical paradigm shift, these areas need to be addressed.  
  
It is this reviewer's opinion that the manuscript would be more defensible and better received in the wider community if gaps and potential concerns are addressed. To wit:  
  
1. To some, the "democratization of [any professional field]" can be viewed, alternatively, as a "shackling of the [professional]" to the lowest common denominator. In the case of a scientist, for example,  
a. When the scientist is required to work in a fishbowl and defend (or at least "sell") every step of his or her work, the process becomes a beauty contest or a sales competition, rather than a scientific pursuit. The very reason we have civil service is to provide a measure of stability and impartiality, and insulate professionals to some degree from political biases and public whims.  
b. Subjecting the processes (rather than merely the inputs and outputs) of a scientist's work to the political will of nonprofessionals denies them their professionalism and erodes the very idea of what it means to be an educated, knowledgeable specialist. The silent corollary is that every aspect of a professional's work should be made publicly understandable, so that public judgment can replace professional judgment in the course of work. Yet there is a reason that our society is built on Division of Labor - it allows professionals the luxury of mastering special knowledge and understanding that not everyone else has. This specialization advances the society as a whole.  
c. An example may be illustrative. My doctor and I are user (manager) and consumer (stakeholder) of what the pharmacist dispenses to me based on a prescription, and we both have high stakes in the quality of the pharmaceutical products. We may express our political, personal, or professional will regarding what drugs are brought to market, and we may express our (dis)satisfaction with the end results. However, we have no expectation that we be directly involved in the pharmaceutical development process. Rather, we trust that the work is peer reviewed, tested, and judged by the appropriate professionals whose special knowledge is appropriate to the task.   
  
2. The authors quickly dismiss the idea of any government agency or personal "ownership" of the data, methods, work processes, and work products of government scientists. They also characterize bioassessment as different from other science in that it is meant not to stand alone (for example, simply to determine whether a waterbody is healthy and supporting aquatic life), but to serve the aims of managers. The unspoken corollary to this view is that bioassessment scientists are not "real" scientists, but rather mere functionaries. I would caution the authors to be careful here. On a regular basis, many government employees design protocols, methods, and studies, and then generate data, tools, reports, and publications that meet or exceed the standards of academic work. Moreover, their work products are not biased by a "positive results" publication imperative, and their data have the added advantage of residing in an institutional repository where they will be retrievable for generations, or as long as the internet infrastructure is intact.  
  
3. We live in an era in which many loud political voices devalue science and even denigrate and attack it. Many stakeholders are, in fact, critics who have particular self interested motives and will find any chink in the armor of those who seek to protect natural resources. The authors seem to tacitly assume that opening up the workspaces of professional scientists to the public and to policymakers will not make our natural resources more vulnerable, but rather will restore public trust in our work and our institutions. This belief seems, at best, naïve. It would be helpful if the authors could explain the mechanism by which the shrill enemies of science would become allies, if only they could impose their will on the process - and how this would not erode the foundations of impartiality upon which the scientific method is founded.  
  
4. The authors have demonstrated clarity of thought and expression. It is very good that they draw a clear distinction between "open data" and "open tools," versus "open science." It is true that government could do somewhat better with "open tools," though more and more government scientists are converting tools to the R programming environment and contributing packages to CRAN, for example. However, the authors do a significant disservice to government agencies by failing to catalog or even mention the extensive effort that has been invested in ensuring both the quality and availability of data. Some of these include the following:  
a. Data sharing and data comparability are a centerpiece of federal water monitoring programs, under ACWI and the NWQMC. https://acwi.gov/methods/pubs/over\_pubs/valcomp\_fs.pdf   
b. Federal agencies are already committed to open data. USGS (which operates NWIS and BIODATA) and USEPA (which operates WQX/STORET) collaborate via the Water Quality Portal and the Internet of Water, as well as maintenance of the National Hydrography Dataset and its derivative products, in order to make high quality data available to the scientific community and the general public.   
c. State and Tribal data collected under EPA funding is required to be uploaded to STORET/WQX.  
d. (Side note - the SCAPE model described in this article appears to have been built on StreamCat, an EPA produced, publicly available geospatial dataset. The SCAPE website clearly credits StreamCat; this manuscript probably should as well.)  
  
5. Although some thresholds are predictive (for example, 30% impermeable surface), landscape based watershed models show relatively poor correlation to macroinvertebrate assemblage quality, whereas instream habitat characteristics, certain water quality chemistry measures, and riparian corridor characteristics have better predictive value. The authors of the SCAPE model undoubtedly understand this as well as the nuances of natural stream classes, metric development, etc. Is it fair to ask, however, whether even the most thoughtfully designed tool can be misued by someone who does not understand it? Is it possible that the SCAPE tool could lead the novice/layperson user to conclusions or decisions that are distorted or entirely inappropriate? It is a relevant question.  
  
6. This manuscript focuses on bioassessment data in particular, but it gives scant attention to the particularly thorny nature of biological data. The authors cite Cao and Hawkins 2011, but they do so in a general discussion of "duplicated effort" and "lack of coordination in the monitoring community." The same paper, in fact, is a good review of the particular difficulties of biological data. This manuscript would benefit from a short discussion of the unique characteristics and problems of biological data that set it apart from other data types, such as nested hierarchy, changing taxonomy, ambiguous taxa, the importance of ancillary information such as ecological and toxicological data, etc. Other papers in this vein are Cuffney et al 2007 (JNABS 26:286), Lenat & Resh 2001 (JNABS 20:287), Chessman et al 2007 (JNABS 26:546) and Stribling 2011 (Chapter 4 in the book "Modern Approaches to Quality Control.") 

**Reviewer 3 comments:**

Summary: In this article the authors outline the ways in which open science principles and technologies can be applied to bioassessment to better link science and management. This is done by summarizing open science core concepts, and new and emerging tools for application by researchers and managers. It is further illustrated by describing recent development of a bioassessment products using an open science approach.   
  
General Notes: The topic is important, and the article is generally well written. The primary drawbacks were that it seems much longer than is necessary, is repetitive, and does not always make clear how open science approaches will advance bioassessment. It can be "preachy" at times, assuming that end users are very reluctant to adopt these approaches instead of (as I suspect) mainly overburdened and lacking appropriate support. For example, most research projects have 1-2 years of funding; management agencies and their priorities are funded on 2-4 year legislative cycles. In this context, producing a grey literature or peer-reviewed article may be the only/most viable option for communication. In this same vein, who is the target audience(s) - is this article intended to convince funding sponsors and/or legislators holding the purse strings? State bioassessment program mangers? Academic researchers? In my experience, the hurdles may differ among groups. For example, academic researchers may be in a better position to adopt new tools, but have less access to manager/stakeholder input or ability to sustain an interactive website over time. Conversely, state agencies may be in a position to sustain a website over time (maybe!), but have less bandwidth to explore and learn new tools given their management mandates. Collaborations may be an important and critical way to overcome these kinds of institutional limitations for open science.  
  
The article would be stronger if it were shorter and focused less on convincing the reader, and more on encouraging/hand-holding by 1) demonstrating the benefits 2) frankly acknowledging challenges, and 3) presenting open science as a series of "components" that facilitate a more transparent, repeatable, and iterative/engaging process (vs. assessment itself being "open" or "closed').  
  
The specific comments below all stem from the general notes.  
  
Introduction: Is too long. It has good information, but makes the point in multiple ways that bioassessment is uniquely embedded in legislation/management action and could benefit from adoption of open science. For example, the exact words "hundreds of assessment methods" appear in the 1st and 3rd paragraphs, and similar points appear throughout. The challenges presented are accurate, but I have a hard time seeing how adopting open science will necessarily meet all of them. For example, abundance of methods and lack of guidance (Lines 90-93) is a problem, but having these same methods available more openly wouldn't necessarily provide guidance on which ones are most appropriate. Also, the way that challenges are interspersed with how open science provides solutions are muddled and difficult to follow. For example, "Biological indices are typically used to develop post-hoc diagnoses to trigger remediation or restoration actions, or to serve as early warning indicators of environmental change" (Lines 63-64) - but then "discoverability" (Line 67) is presented as the solution to this challenge.   
  
I think a better approach is to briefly summarize the benefits of open science (either in bioassessment or other fields) and present it as one way to help address the science-management gap that is currently identified in bioassessment, without trying to map out how it will exactly address very specific challenges.   
  
Line 56-58: What is "imbalance" here? Seems to need a citation too  
  
Lines 108-123: Repeats much of the abstract. Recommend writing this as objectives, and - more importantly - identifying your audience(s) and what they should get from reading the article.  
  
Lines 97-103. I wouldn't call out that bioassessment has or has not embraced open science compared to other disciplines; it's very subjective - and constraints of time, funding, and expertise are the likely culprits. Sustaining management attention can be a big hurdle too for agencies. However, there is a case that bioassessment could benefit somewhat uniquely from these tools, due to the nature of being embedded in legislative mandate/public interest (requiring transparency), being a relatively new mandate/concept (requiring ongoing development), and the need for each state to conduct and report on monitoring (requiring replicability).   
  
Lines 125-153. I don't think it's necessary to elaborate on the distinction between "conventional" and "open-science" approaches - managers and researchers are pretty aware of the limitations of current approaches; besides it's already covered in the Introduction. Defining what you mean by open science here should be sufficient, without calling out everything that's wrong with non-open science.  
  
Line 154-201: Why is open data not part of the open science principles section? This section also seems very long to make a point that is widely accepted (although one that can be difficult to implement and - more importantly - sustain)  
  
Sections "Applying open science principles to bioassessment" and "Conventional bioassessment" are making the same points, but from an inverted perspective. Example: Lines 255-256 is about accessible data and lines 272-273 are about inaccessible data. Also, lines 261-263 state: "A typical workflow for developing a bioassessment product is not entirely dissimilar from a conventional scientific process" - exactly. So it's not necessary to elaborate on the standard process, it's the one everyone knows. These two sections also repeat the same challenges raised elsewhere.   
*Agreed that there was some redundancy between these sections. We have revised the “Applying open science principles to bioassessment section” to serve only as an introduction to the following sections. Also, please see our response to the AE on the same topic (i.e., response to comments on lines 202-258.*

Lines 296-297: they are uncommon in all disciplines - it doesn't seem helpful to call out the bioassessment community

*This sentence was revised to place a more positive emphasis on inclusion of these tools, rather than their current exclusion from bioassessment. We feel this creates a more pragmatic description. “…from conception to completion (e.g., for environmental flows,* [*https://eflows.ucdavis.edu/*](https://eflows.ucdavis.edu/)*). Similar approaches could be useful in bioassessment to remediate data accessibility and reproducibility issues.”*  
  
Lines 309-314: This is fairly redundant with Figure 4 (but doesn't reference it). Also is related to the sections after, but not aligned with them.   
  
Lines 309-492: (Related to above) Coming up with a more streamlined way to refer to the recommended process will help organize the paper, avoid redundancy, and help readers follow along. In other words, I'm not sure what the intended organization is here and in the following sections - is it organized by process or by tools? Make use of the tables and figures whenever possible instead of walking readers through text-based descriptions of tools/packages. Also, I would probably avoid sticking in more justifications for open science approaches (e.g., lines 348-351) in these sections, since (hopefully) the reader has already bought in to the "why" at this point.  
  
Lines 364-492: This section is too long, and walking readers through Table 2. If needed, expand Table 2 with some additional details or columns (e.g., example applications, which bioassessment steps it will help with) and dramatically cut down the text.  
  
Lines 459-468. Suggesting readers can start developing R packages for bioassessment is an off-putting stretch/tangent - if they are able to develop R packages, then they are already experienced in using these tools.

Line 493: Suggest subtitle "Open Science in Practice: The SCAPE Project". In this section, be sure to refine and present information that is relevant to open science and/or is not redundant with previous sections. For example, lines 526-537 rehash challenges and the science-management disconnect. Similarly, struggling to see how lines 538-552 relate to open science.

*Subtitle was changed as suggested. Much of this section was also reduced following recommendations herein and from the AE to reduce redundancy.*   
  
Lines 588-631: I refer to this later (Figure 4), but this section is essentially a repackaging of Table 2 and Figure 3. Recommend cutting this and/or finding ways to integrate anything really essential or new into other sections.  
  
Section "Limitations and Opportunities". This is an important section, and the abstract states: "We also discuss technical, sociocultural, and institutional challenges for adopting open science, including practical approaches for overcoming these hurdles in bioassessment applications.". But I didn't find many practical suggestions for overcoming these hurdles in this section. For example, "Many scientists feel they cannot prioritize learning new skills given existing demands on their time" (648-649) and "requires a research team to stay abreast of new technologies as they are developed" (652-653)are big hurdles that won't be overcome by a better appreciation of the benefits of open science. Many of these hurdles are based in overarching constraints on the way science and management are funded and sustained (at the whim of short-term funding and even political cycles). My reading of this section is that it goes after the low-hanging fruit (e.g., scientific culture being closed to new ideas) instead of offering suggestions to counter the more intrinsic hurdles of lack of sustained funding, personnel, and expertise to develop open science models. I think focusing on suggestions that help overcome these types of hurdles will be better received and more reflective of the current challenges faced by research and management groups.  
  
Table and Figures  
  
Figure 2. The "a" section detracts from this figure. I can think of many "conventional" assessments that incorporate some aspect(s) of an open science approach, but fall short of incorporating all aspects of it, due to lack of time/expertise/money etc. This figure would be better presented as "Idealized or potential components of bioassessment based on open science principles" to encourage adoption of different components - piecemeal if needed - vs. an all-or-nothing buyin.  
  
Figure 4. This type of information is already presented in Table 2 and the text (and to some extent the workflow figure previous) - the large graphical representation seems unnecessary. Some information could be added to Table 2 if needed to replace this (i.e., more explicit explanations or examples of R package applications).