Dear Dr. Bonte,

Thank you for these reviews and your candor on how to address them. My apologies that this was a lengthy revision process, but I've been in the field or teaching field classes non-stop since I received these followed by a delightful family emergency. As such, this is marked as a new submission, but was previously submitted in the system as OIK-09402.R1, which you had assessed as Minor revisions.

With respect to your comments, they are completely reasonable, and I have addressed them and additional feedback from reviewers as requested. Below, please find responses to reviews where appropriate. In the attached manuscript, I will highlight where each comment has been addressed as well. I look forward to any further thoughts! Please note, I have already received several inquiries regarding the preprint version of this manuscript (as the technique has remained unchanged - merely the text surrounding it) and it seems to be already making its way into the multifunctionality literature! I look forward to Oikos being the vehicle for this advance in the field!

- Jarrett Byrnes

Reviewer(s)' Comments to Author:
Reviewer: 1

Comments to the Author
The authors have now revised the text and provided point-by-point responses to my earlier suggestions and questions. The additions (e.g. new paragraph in the discussion) and clarifications in the text have further improved the manuscript. The new table 1 nicely summarizes the different choices that need to made when calculating the metric(s). So, overall, I am definitely satisfied with the way the authors have dealt with my feedback. I also had a look at the comments and responses of the other referee (reviewer 2). Taken together, the rebuttal letter provides an interesting exchange of ideas, several aspects of which also ended up in the manuscript and are therefore available to readers.

We are delighted with your take on this manuscript! Making this table really helped in making the manuscript more accessible to a broader audience. Thank you!

I have looked closely at the table 1 and have the following suggestions for the authors to consider:

R1Q1: 1/ The first column says 'parameter', while this column is actually presenting different types of variables (derived from empirical data, user-defined constants, derived metrics…). Not

sure what other term is more appropriate here. Just 'variables'? A more specific term could be added to each variable to clarify what it does in the derivation of the metrics.

We have changed the name of the column and edited the table 1 caption accordingly. We have also added an additional column to clarify the nature of the variables, as well as reordered the table.

R1Q2: 2/ The description of q mentions 'larger values upweight the contribution of evenness'. I think it is clearer to say that q is expressing the degree to which high-performing functions are upweighted?

This is an excellent and much better way of putting it. We have altered the text accordingly.

R1Q3: 3/ The description of tau says that it is threshold for functions to be considered independent in the calculation of ^qN. Shouldn't this refer to the correlation-corrected effective number of functions, so ^qN_u ?

Great catch. Thank you!

Reviewer: 2

Comments to the Author
*The revised version of the article is improved in some areas, though not all my comments were addressed in full. The authors have instead opted to keep the article short and in a similar format as before. I think this is ok, and I still like the paper, though I do feel that the authors stated aim of setting a new standard is unlikely to be met unless some of these issues are resolved. As a result, the paper, and its significance, does seem a bit overstated in its current form, from my view at least. In particular, the biological aspects are somewhat glossed over, but maybe this is best addressed in another paper? As stated in my previous review I believe its these issues that I think most researchers outside the BEF world have a problem with, rather than the mathematics of the metrics, so perhaps the mathematical resolution should be stated as the advance here, with more explicit acknowledgement of what still needs to resolved elsewhere. What is still missing a clear description of what *biological* insight the metric gives that other existing metrics do not- maybe that could be added at the very least, or an acknowledgement of the cases when they provide similar information? I present some additional, related, suggestions below.*

We are delighted that this manuscript has produced a spirited back and forth with R2. We also see that R2 is struggling with many of the issues we ourselves struggled with in writing the paper to craft it for the intended audience (practitioners seeking to get a better understanding of what multifunctionality is and how to measure it - from grad students to more senior scientists). It's no easy feat, and we are glad to have Oikos's Forum section for this piece. Some of the wider questions that the reviewer brings up could be the basis of articles of their own. We

therefore prefer not to expand too much on them here and instead concentrate on quantifying multifunctionality.

We have sought to use the reviewer's specific comments as a guide to answering their concerns, and hope we have achieved this sufficiently for the editor.

Specific comments:
R2Q1: 4 multifunctionality metrics measure (mf is a concept beyond being a measure)
Fixed

R2Q2: 27- say that that is what it is defined as here- as it's really not that standardized in the literature or the minds of researchers, especially beyond the BEF field

We have made the sentence more specific. We note that, for the past decade, standardizing this definition is precisely what we, the authors, have been seeking to do using this definition (which is how MF is used colloquially throughout the literature).

R2Q3: 29 biodiversity – ecosystem

Sorry, we are unclear as to what the reviewer is asking. That we add a - between biodiversity and ecosystem in "biodiversity ecosystem function"? If so, unless the editor deems otherwise, that is incorrect. See the titles of books on Biodiversity Ecosystem Function relationships as well as many reviews that do not use a hyphen.

R2Q4: 31- no, the concept was already there in land management long ago – it is the BEF metrics that have spread recently

We have added additional language to this effect.

R2Q5: 33-34- I think it would be useful to state here what ecosystem multifunctionality is in this paper, if the wider discussion of biological meaning is not to be addressed.
This is addressed at the start of the paragraph where we explicitly define multifunctionality.

R2Q6: 41- check -'and this can'?
Sentence modified.

R2Q7: 49- where high multifunctionality
Fixed.

R2Q8: 66- this is one of several definitions of multifunctionality, so perhaps state as which 'we define here as….' Or 'the definition…. widely used in BEF research'

Bringing together the definitions into something unified was, indeed, the point of Byrnes et al. 2016. Hence it's citation, and we are clear that we are using the definition from that manuscript as you had previously requested.

R2Q9: 147-153- I find this section, which is clearly key, hard to follow- what does this mean biologically? In what research would we want to upweight and down weight functions just because they are at high or low levels?

We have re-written the end of this paragraph to be clearer, and also included some clarifying material in the next paragraph. As with species diversity, it depends on the researcher or manager's desire to value presence/absence versus relative abundance/functional performance. See Lines 152-154.

R2Q10: 166- why does a manager want a system that has high levels of functions, i.e. a 'fast system' (as opposed to services) or with a few randomly selected/arbitrary ones at very high levels? I don't see how this can inform management (see earlier comments from the last round) so maybe remove the statement about management here.

We disagree. Some systems might be most useful when a small handful, but not all, functions are performing at high levels - in particular if the identity of those functions can differ from year to year (i.e., the identity of different fisheries stocks over time, not all of which need to have high yields in any one year).

R2Q11: 222- this assumes that high function = good- it may not be – there are good and bad levels of services but not of functions – please revise wording accordingly i.e. well>at high levels

Great point. We have made the revision.

R2Q12:235- services now pop up out the blue so perhaps say something like 'when this approach is applied to ecosystem services'

Great point. We have made this change.

R2Q13: 236-237 – I think the issue is broader than this example – see also greenhouse gas and water fluxes and how these relate to other process rates.

We agree. We hope the example provided will get people thinking more about flux/storage tradeoffs.

R2Q14: 257-259- I actually think this new discussion of ES bundles muddies things further here as the revised paper has now moved to concentrate on functions, which helps focus the paper, and how we deal with services is another matter that requires different approaches, and a different paper (I don't think the proposed metric is particularly suitable for services). I would suggest removing this line.

We disagree. This section maintains the connections between two linked literatures - and the methods are certainly adaptable to the ES literature. There is a long history of back and forth, as we highlight based on your suggestions from the last review, and feel this adds a really nice additional dimension to the paper - and more opportunities for future crossover and collaboration. In particular, it links to an approach used in the ES literature to deal with the same problem, giving many of our readers an opportunity to dive deeper into that literature themselves.

R2Q15: 322- I'm sorry, but I don't find these metrics particularly intuitive as the paper stands

We agree that this new framework requires some additional thought. As such, we have softened the language in this sentence.

R2Q16: 330-342- these are more about the measurement of multifunctionality itself than what the application of this method would tell us about the functioning of ecosystems

While many of these areas are directly related to measurement, we wrote this section to emphasize the power of thinking carefully about this metric to answer basic ecological questions. We strongly feel - and indeed this is to some extent why we wrote this manuscript - that with the proper metrics (and development of theory around them) in hand, we can open up new areas of research. Without a ruler, one cannot measure the height of a tree canopy.

R2Q17: 353- what manager wants lots of functions operating at 50% of their observed maximum or more? I really think discussion of management should be omitted here

We agree 50% for all functions is likely not ideal. However, in many scenarios, different thresholds can be useful for different functions. As two authors here are marine, we often think in terms of zoning and multi-user coastal areas, where thresholds for different ecosystem functions and services (wind and wave power, fisheries yields, optimal bycatch reduction, whale entanglements, etc.) comprise a diverse portfolio which have very different thresholds for each function or service. Indeed, one author (Byrnes) has had conversations with environmental studies PIs about using thresholds as a potential metric for these systems. As such, we have chosen to leave it.

R2Q18: 368- delete comma after idea?
Done