

# Response to Reviewers for Understandable Multifunctionality Measures Using Hill Numbers

Dear Dr. Bonte and reviewers,

We thank you for your reviews and edits. They have been most helpful in crafting what I hope you will find is a stronger manuscript for publication in *Oikos*. Below, we answer all comments point-by-point with reference to line numbers in the revised manuscript. We have also included a version of the manuscript with track changes and a clean one with labels highlighting changed areas with specific reference to reviewer comments (with tags e.g., R1 Q1).

Thank you for your consideration.

Sincerely,

Jarrett Byrnes, Fabian Roger, Robert Bagchi

Recommendation by the Subject Editor (Prof. Dries Bonte):

As you will see, I received two excellent evaluations for your manuscript. Quantifying multifunctionality is indeed a timely (and maybe highly needed) objective in the field of BDEF. While I do share the view with rev2 that a mathematical framework is maybe not the most urgent advance in the field, I do see its value to proceed in general. The reviewers give good suggestions on the presentation and framing of the metrics, and on the need to discuss (a bit) into more detail how the suggested approach can/needs to be combined with sensible biological interpretation and rationale of multifunctionality. I also feel the latter is important. So, overall, I like this manuscript, some revisions are needed and I am looking forward to see the new version. A convincing revision will likely not be sent back to reviewers.

**Thank you for the kind assessment. We agree with all of your points, and will address them point by point with respect to reviewer comments below.**

Reviewer(s)' Comments to Author:

Reviewer: 1

Comments to the Author

Although the concept of multifunctionality has been used in the literature for some time now, the development and use of metrics has not yet been fully consolidated. In this article, the authors build on a solid theoretical framework (cf. species diversity) to propose a new approach. It integrates various ideas on multifunctionality indices that emerged in previous studies. And the proposed approach also provides a simple and generalizable set of metrics that can be further refined based on further theoretical research.

**We are delighted that R1 sees why we wrote this manuscript and is on board. Thank you!**

The manuscript is well organized, clearly written and convincingly positions the new approach in the current state of the art in this research field. I have two more global comments and several smaller suggestions that I hope will help the authors to further improve the text.

The first general remark is that, throughout the text, the authors seem to suggest that the approach they propose, as a unifying concept, is the main way forward in this domain (e.g. lines 34-41 and lines 280-282). There are indeed reasons to assume this, as is clearly explained in several places. However, a reader may also read the text as if it is proposed that existing metrics become redundant. I would emphasize more strongly that the proposed framework can give a boost to the development of theory and empirical research on what drives multifunctionality (e.g. lines 293-295), without necessarily replacing existing metrics altogether. They can be complementary. I am convinced that existing concepts (cf. 'the current state' paragraph) continue to have their place because:

(1) They are also consistent with the definition of multifunctionality

(2) The use for one or the other metric depends on the context of a research project. If one is specifically interested in which functions exhibit trade-offs or synergies, a multivariate approach (e.g. Dooley et al. 2015) where covariances are estimated may be very relevant. In a highly applied context, e.g. where the desired level of functioning is known a priori, a threshold approach provides a simple and easy to communicate metric (cf. lines 202-204). The latter is important for practitioners.

**R1Q1. Ah, it was not our intention to make it look like we were invalidating previous work! Far from it (and you can see results from our analysis of the Duffy et al. 2003 work show broadly similar answers). To make this clearer, we have added a paragraph in the discussion on lines 344-360 to discuss how previous metrics were related to or a subset of this approach, or sought to answer a wholly different question, and as such are complementary.**

As a second, more practical suggestion for improvement, I suggest that the authors consider making an overview (e.g. table) of the various choices to be made when calculating an index in this framework. There are obvious choices, such as the order 'q' of the index and the way in which data are standardized. Other choices are less obvious, such as the choice of a distance matrix and the associated value for 'tau'. These different 'user choices' are now scattered throughout the article. By bringing them together in an overview, the text becomes more like a roadmap. I think this can further facilitate the use of the framework. Note that this overview can also indicate which aspects of 'user choices' require further research versus which are +/- established.

**R1Q2. This is a fantastic suggestion! Indeed, we've wrestled with q and tau as well! A succinct table is a great idea and we have added one (Table 1)**

Smaller suggestions.

- Does the new framework have the existing approaches (cf. section 'current state of...') as special cases? This aspect has been clearly discussed in the context of species diversity metrics. The text clearly mentions when the new metric converges to the simple averaging approach. What about the other (e.g. threshold approach)? Or are they just very different in terms of what aspects of multifunctionality they express (and are thus complementary, cf. first main comment). Would be good to briefly elaborate on this in the text.

**R1 Q3. Yes,  $q=0$  is the averaging approach. See the new paragraph in discussion mentioned in response to your larger comment above. Thanks for the suggestion - it's a great way of making a connection to previous work.**

- The generalizability to other fields (e.g. line 45 and 193-195) can be explained in some additional lines. I found it a bit vague right now.

**R1Q4. Our intention is to show that this metric can be used as any other univariate metrics in different fields - including whatever intellectual traditions and forms of analysis those fields use. We have added this concept to lines 227-229**

- Line 94: "we have noticed that those who do often report a single threshold in their main text anyway". Did you quantify the number/proportion of studies in which this was the case?

**R1Q5. We did not quantify this, nor do we wish to call out individuals, as this would be a distraction (in fairness, one of us is referencing their own graduate student's masters work before they became a PhD student in their lab!). As such we have changed the text to "many of those..." and incorporated that even then, full explorations are often present - just in supplements. See lines 106-108.**

- Line 132: I think the explanation of what the effective number of functions means in general terms can be clearer. For readers who don't know the rationale behind Hill numbers in a species diversity context, the text in *italic* might not be sufficient.

**R1Q6. No problem! Indeed, after submission, we had a conversation with a colleague who read the preprint and, in the course of the exchange, wrote some text to which the colleague replied, "Yeah, you should have put THAT in the paper." We have thus added text to lines 149-153 that you and our colleague (Jon Lefcheck) thought clarified the meaning of effective number of functions and our multifunctionality metric. See also lines 175-175, 182-184, and 194-195.**

- Equations (4) and (5): the symbols of the indices can be explained in the text. Why " $qM_e$ "?

**R1Q7. Fixed on lines 187 and 191.**

- Line 170: "Why must we consider the level of function and effective number of functions in one metric?" The authors provide three types of reasons: mathematical (lines 171–180), conceptual/definition related (line 187 – 191), and convenience (line 192-195). Personally, I would start with the conceptual reason.

**R1Q8. As explained in response to R1Q6 above, we have added further justification in multiple places.**

- Figure 1: The x-axis corresponds to the arithmetic mean functioning, i.e. variable A? In general, the caption could be clearer about which variables are shown and the symbols that are used in the main text can be added to the axes titles.

**R1Q9. We have revised the text of the caption to make this clearer.**

- The new framework has been integrated in the multifunc R-package. This info is hidden in the appendix. I would also provide this info in the main text and, if the authors decide to provide a summary (table) of user choices, also refer to the related syntax.

**We thank the reviewer for pointing this out. Definite oversight on our part to do the work and then not point it out! We have added reference to both in the discussion paragraph mentioned in R1Q1.**

Reviewer: 2

Comments to the Author

This is a well written article that presents a new metric for measuring ecosystem multifunctionality. The proposed metric is based on existing measures for measuring alpha diversity. The article is a good read- enjoyable and thought provoking, and a good match for Oikos, which I've always seen as being an 'ideas' journal. I enjoyed debating with it - agreeing with some points but not all- while reading.

**Thank you! This is exactly the kind of read we were hoping to give people in the field!**

While I like the article and think it makes a useful contribution to the literature on this topic, I also think it somewhat misrepresents the biggest issues with the ecosystem multifunctionality concept and its measurement. As a result, I'm not sure the paper resolves as much as I think it would like to. Nevertheless, it is interesting and should stimulate further debate, and hopefully encourage readers to think a bit more carefully about what these metrics represent. I would hope the suggested changes below will push this further still.

The paper starts its argument from a position in which the biggest issues with multifunctionality measures are mathematical. Clearly this is the authors view that they are entitled to, but I don't

think these are the biggest problems. It advances on the mathematical issues, but more fundamental biological ones remain, and this should be acknowledged.

There is a deeper problem with defining what we are really trying to measure with these metrics and what a system that has high process rates for multiple functions represents. If we could agree what we really want to measure, the correct metric will present itself. In the proposed metric and most those currently existing, multifunctionality is high if many functions are at high levels at the same time and in the same place. Such an 'all systems go' ecosystem with high process rates, which could also be described as a 'fast' system (assuming we don't treat stocks as functions) would score highly on the current metric and most of the others out there. Why is that 'good' or why does it indicate a desirable property? Alternatively, if we don't include such value judgements then how should we view a high multifunctionality ecosystem from a biological perspective? For example, a peatland would be a low multifunctionality system in the proposed metric, compared to a fertile grassland (on a drained peatland for example), as its process rates are low, while at the level of ecosystem services the two systems would provide different and contrasting bundles of services. A full discussion of these issues is clearly not the scope of this paper, but I think it's important to at least acknowledge them and to discuss the \*biological\*, not mathematical, meaning of the proposed metric and how to interpret it.

**R2Q1. We are delighted the reviewer is engaging with the material this way! We disagree only slightly here. We actually think one of the core problems of the field is that, while we agreed what we want to measure - the simultaneous performance of multiple ecosystem functions (see Byrnes et al. 2016 which formalizes this and draws on previous similar statements from other papers - and we have added this definition to lines 66-67), we have not been able to arrive at a single metric that does so in a way that all parties involved feel is appropriate. Would that have been otherwise, that 2016 paper might have been the last word. We have added text regarding how multiple metrics have gotten close, but that the one presented here actually achieves the aims in the discussion. See our response to R1Q1. We agree that we needed to make the biological meaning of the metric clearer so that it more cleanly matches up with the idea of total-system measurement. We have done so in our response to R1Q6.**

Some additional and related points that I feel should be covered are:

What biological insight would be provided by the suggested metric that other metrics do not give? E.g. an average and variance? If the proposed metrics are similar and highly correlated with other existing measures (as suggested by the appendix results) is there a need for them?

**R2Q2. We are glad that you and R1 were on the same page! See our response to R1Q1.**

The suggested provides two metrics – the question for me is what do these represent ecologically and why should we care if an ecosystem is high or low in these properties (especially beyond the BEF more species = more functions argument)? Alternatively, if the metric is BEF research oriented it should acknowledge this.

**R2Q3. We hope that our response to R1Q6 addresses ecological meaning more clearly (again, that you both asked these questions means we needed to clarify!). With regards to the ‘value’ of high versus low measurements - we feel that is beyond the scope of this paper and would cloud our focus too much to address.**

It would be nice to see this method applied to some real data so we can see the metrics and how to interpret them – this is partly provided by the appendix- maybe some of that material can move to the main paper?

**R2Q4. We realize our error was to not include direct reference to the supplement. We have done so in lines 350. Unless the editor disagrees, we would like to keep the worked example in the supplement in order to maintain brevity and focus. We are open to changing this if the editor so desires, however.**

The comparison with species diversity measures is interesting but species are definable discrete entities, functions are not (relatively speaking of course). This is covered in part in the correlations section but there is an issue in that if only mathematical correlations are used to adjust function weightings then if there are problem is there are biological reasons for the correlation- e.g. all functions are coupled even if quite different conceptually- often the case for carbon, nitrogen and water, and/or responding to a common driver. Also, mathematically, can a system be high in multiple uncorrelated properties- i.e. does correcting for correlation cap the multifunctionality that is attainable? These issues should be discussed.

**R2Q5. We agree that the correlation matrix is not always correct biologically. We have added text to this effect on lines 284-286. In our text, we present several different approaches, but a set standard is beyond the scope of our current paper.**

Specific comments and suggestions:

4- provision indicates that functions are services- are functions ‘provided’? performance, as used on line 25, is a less loaded term.

**R2Q6. We have made this change on line 27 and in the abstract.**

26-29 this misrepresents the literature- multifunctionality as a concept has a much longer history in land management/ecosystem services etc than it does in the BEF field. It’s just that the BEF field pioneered the metrics, and these spread into other areas of ecology where the idea of multifunctionality already existed- see e.g. Holting on this.

**R2Q7. We agree, and as such have added additional references to this literature throughout the text and delineated community ecology versus science in general in lines 28-30, 36-37, and 257-259. We have also included two additional references to this literature that were included in earlier versions, but cut in edits. We thank the reviewer for pointing out this omission.**

44- I feel this is the bigger problem- a lack of discussion as to what we actually mean by multifunctionality- once that is decided the metric will present itself.

**R2Q8. As mentioned above, we have inserted the definition back into the manuscript in order to strengthen our argument regarding the Hill Number approach.**

45-46- multifunctionality is widely used outside the BEF field now, and as been for a long time.

**R2Q9. We agree. We have included more on extending it outside of BEF in response to R1Q8.**

57- a point of discussion here is that if these advanced species diversity metrics are available why is species richness still the most used? This is something to consider when suggesting similar for multifunctionality.

**R2Q10. Whether richness is the most used metric really depends on subfield. However, a discussion of the use and value of richness versus other diversity metrics in Ecology is beyond the scope of this MS.**

83- at or above

**R2Q11. Fixed. See line 94**

103-104- I read this a couple of times to get the point and so would suggest a laymans term version next to it. I guess what it is, sort of, saying is that it's a gradient from a pure richness count to a heavily weighted towards evenness measure, with various weightings of these elements in between. Or maybe I misunderstand?

**R2Q12. Yes, that is it. See also other changes we made for R1 referenced above where we have expanded on this..**

142- please say when you would want different weightings

**R2Q13. We have built a table to describe this more. See Table 1.**

202- this mixes functions and services- if a certain level of a function is wanted it is a service (or indicator of one). We should be explicit when we are assigning 'value' to functions.

**R2Q14. We agree we could have kept the two separate better. We have done so with revised lines 238-239**

244- again, here the problem is as much, or more, biological than mathematical- a lot of functions could be/are correlated, and is it possible to find a system with high levels of several uncorrelated properties?! (hence the slope weakening as this is incorporated) Eg in the BEF experiments almost everything is positively driven to biodiversity, and so correlated, so a single \*function cluster\* may come u mathematically, and the others are unimportant – this would weaken the BEF relationship as it has done here.

**R2Q15. This echoes concerns from above. We hope our revision from R2Q5 addresses this issue.**

278- offering

**R2Q16. Fixed. See line 324**

281- I don't think the literature has grown slowly- there are loads of papers on multifunctionality (though a much smaller number of genuinely insightful ones!) There are also a lot of highly cited multifunctionality papers, which indicates a broad interest, so I feel this section builds something of a straw man. Maybe it should call for more considered studies instead?

**R2Q17. We have qualified this someone on lines 326-327.**

285- I wonder if it would be interesting to see when different alpha diversity metrics are used and seen as appropriate and relate this to when different multifunctionality metric might be used? Given different study objectives is it possible to have a single \*correct\* multifunctionality metric?

**R2Q18. We agree this is an interesting question, but, do not propose to suggest that there is one true value of  $q$  and  $\tau$  to bind them all. Indeed, we talk about just such a need for research into the issue in this paragraph.**

294- jumps to services here

**R2Q19. We have clarified on line 341.**

Thanks for providing the code and walkthrough in the appendix.

**Glad you enjoyed it!**