Point-by-point responses to edits/comments from Biogeosciences’ Handling Associate Editor, Tyler Cyronak on 15 January version of MS “Ideas and Perspectives: When ocean acidification experiments are not the same, reproducibility is not tested” (Williamson et al.)

**Edit/Comment:** Line 24. Delete ‘primarily’ and add ‘currently’ to sentence providing definition of ocean acidification. “I would remove this [‘primarily’]. I think it is pretty well established in the community that what we call ‘ocean acidification’ refers to the reduction in pH caused by anthropogenic CO2 emissions, while any other processes affecting pH are something different. In order to avoid any confusion with past geological changes that may have been caused by something other than CO2, I would add “currently” in front of “caused”.

**Response:** The original wording was closely based on the 2014 IPCC definition of ocean acidification ([AR5 WGII Glossary](#)). Rather than trying to make fine-tune changes to that definition, the sentence has now been shortened to: “Ocean acidification involves a reduction in seawater pH (increased hydrogen ion concentration), currently caused by increasing carbon dioxide (CO2) in the atmosphere”.

**Edit/Comment:** Line 29. Replace ‘variable’ by ‘diverse’.

**Response:** Done.

**Edit/Comment:** Lines 36-37. “I am usually not one to bring up my own citations in reviews, but I think this one by Lydia Kapsenberg and myself is particularly fitting here. It discusses the impact of carbonate chemistry variability under the framework of the 3-pronged concept of vulnerability put forth by Vargas et al. 2017. Kapsenberg, L. & Cyronak, T....”

**Response:** Agreed that Kapsenberg & Cyronak (2019) is highly relevant here; citation added.

**Edit/Comment:** Line 48. Replace “and an intended replication” by “with the intention of replicating”.

**Response:** Done.

**Edit/Comment:** Delete “unambiguously”. “Both anonymous reviewers comment on this word choice as being unnecessarily emotive. Speaking from being on the receiving end of a very emotive/non-objective comment in the literature (see Omega Myth discussions) I think it is important to maintain as much objectivity as possible. Overall I think you do a good job of that in this manuscript, while at the same time highlighting the somewhat provocative language...” [remainder of comment not visible].

**Response:** The authors strongly wish to retain “unambiguously”, since it is a factual and objective description. There is nothing inherently wrong (nor emotive) with being unambiguous; instead clarity is a desirable attribute, both in the titles and text of scientific papers, and also more widely. The problem only arises when a clearly-made statement is not well-supported by the evidence. It then becomes misleading.

**Edit/Comment:** Line 54. Insert “either”.


Response: Done

Edit/Comment: Line 55. Delete “the” and “is”.
Response: The purpose of these edits is not fully clear. The sentence has, however, been re-worded to end: “… either flawed or fraudulent, reflecting earlier concerns expressed by Clark et al. (2016) and Clark (2017)”.

Edit/Comment: Line 61. “Weird change in font size from here forward”.
Response: Apologies. This formatting error inadvertently occurred when the Word file was converted to pdf; final version will be more carefully checked.

Edit/Comment: Line 75. Change “remedied” for “addressed”.
Response: Done.

Edit/Comment: Lines 105-106. “I would divide this into 2 sentences to make it easier for the reader”.
Response: Done.

Edit/Comment: Line 108. Delete “all”
Response: Done.

Edit/Comment: line 109. Change “terminological” to “semantic”.
Response: Done.

Edit/Comment: “I wonder how much of this can be linked to observing effects vs “significant” effects. This brought the first figure in this Nature commentary to mind: https://www.nature.com/articles/d41586-019-00857-9…”
Response: Thank you for bringing Amrhein et al (2019) to our attention. The issue of over-interpreting a lack of statistical significance is an important one, and is now covered by a new sentence: “Second, the limitations of statistical analyses need to be recognized: statistically non-significant results do not necessarily mean there is no effect (Amrhein et al., 2019)”.

Edit/Comment: Line 141. “I think they will experience natural variability over many time/space scales beyond just evolution (see Kapsenberg GCB 2019).”
Response: This sentence has been expanded to say “.. evolution and in their habitat diversity (Kapsenberg and Cyronak, 2019)”.

Edit/Comment: Line 141. Change “have found” to “inherently have different vulnerabilities and”
Response: Done.

Edit/Comment: Line 151. “Maybe replace by something like “a body of”
Response: Reworded to “over-riding the consensus of other findings”.

Edit/Comment: Line 156. Delete “best practice”
Response: Done.

Edit/Comment: Line 161. Change “would seem” to “seems”
Response: Done.

Edit/Comment: Line 172. “Wrong word?” [referring to “that”]
Response: Sentence rephrased: “If both groups recognize that there is a..”

Edit/Comment: Re Figure 1. “This is great, should you also put ‘other studies’ in the first panel?”
Response: Additional text added to Figure legend in response: “Whilst ‘other studies’ are also relevant to situation (a), their importance is increased when interpreting results from situation (b)”.

Edit/Comment: Legend to Figure 1. Replace “pairs” by “pair”.
Response: We consider that the plural is correct since there are two pairs, one in (a) and the other in (b).

Edit/Comment: Lines 180-181. “I like this idea”.
Response: No edit required

In addition to the changes discussed above, the authors have:

- Changed “reproducibility” to “repeatability” in the title of the paper, to make it more general
- Introduced the phrase “repeatability” in the text, to cover both reproducibility and replicability (Line 108: “Whilst both relate to the repeatability of outcomes, the test for reproducibility...”)
- Thanked Tyler Cyronak in the Acknowledgements (with those thanked listed in alphabetic order).
- Included in the Reference List the citations for Amrhein et al. (2019) and Kapsenberg & Cyronak (2019), as suggested by Cyronak. Also a new reference (Dahlke et al., 2018) as an example of differential vulnerability of different life-stages to OA impacts, potentially a major reason for replicability ‘failure’ in this case.