

Reviewer #1

We thank Reviewer #1 for the positive and helpful comments. Our response is provided below:

“First, it should be clear where in the artificial sediment the littleneck clams (L. staminea) live. Pore water chemistry was determined at a depth of 8 cm (Line 172) and compared in the results section with samples taken from the overlying water. Since (carbonate chemistry) parameters are different for these two samples, the sediment likely contains a gradient of conditions. It should be clear where the clams exactly lived (even if it varies between specimens, between conditions or in time) to indicate what conditions they experienced. Now it is implied that they all lived at a depth of 8 cm and that those pore water conditions were the ones experienced by the clams.”

Response: Please see the new text starting at line 179. The precise vertical position of the clams was not measured, though they were limited by their small size to the top of the sediment (generally above 8cm). The 8cm depth for pore water measurement was chosen based on the amount of pore water needed to run the analysis, which was ~250 mL per sample (one 125mL sample for pH and one 125 mL sample for alkalinity). Because of the large volume of pore water collected, our measurements represent an average of pore water chemistry, as opposed to a measurement of pore water chemistry at 8cm specifically.

“More importantly, the Discussion lacks the necessary depth. I suggest the authors extend the discussion by comparing their results with earlier experiments with bivalves under OA-conditions. Is reduced growth under acidified conditions reported before and were responses comparable? Or do they seem to vary between species? What does this mean for natural environments in which the sediment is rich/ poor in carbonates? Is the applied level of acidification representative for a near-future scenario? How fast does the shell hash dissolve? And how does this fit on the growing literature on alkalinity enhancement to mitigate acidification?”

Response: We have added three paragraphs to the Discussion to better consider these issues. In short, we conclude that shell hash can buffer acidified sediments and improve conditions for bivalves, but the biological benefits are inconsistent and depend strongly on factors such as species, life stage, predators, vegetation, and local environmental conditions. Overall, shell hash appears most promising as a short-term, site-specific habitat-engineering tool rather than as a broadly effective mitigation strategy.

Reviewer #2

We strongly thank Reviewer #2 for the detailed feedback. In short, we agree that there are better statistical approaches than those used in the original version of this manuscript. We have added a new co-author, Dr. Gabriel Ng, who helped us reanalyze the data using GLMMs, as suggested. The statistics have been completely redone, with new figures, text, and code. Our point-by-point response is provided below:

“The problem here is that pH and header tank are completely confounded. Any difference between ambient and acidified treatments could be due to pH itself, or some other/unmeasured characteristic of the header tank (microbial community, chemistry, etc.). Because there is only one header tank per pH level, there is no way to separate these effects statistically. This is particularly concerning because the seawater in the header tanks was replaced every three weeks (line 153), which may have allowed water quality to evolve differently between tanks and deteriorate independently of pH, thereby confounding the interpretation of treatment effects.”

Response: We agree that the experimental design does not allow us to separate pH from header tank as independent experimental factors. The revised manuscript explicitly acknowledges this limitation (see lines 290-291 in the revised manuscript). We also include paired t-tests on the water chemistry between the two tanks to demonstrate that pH is the only statistically different value measured.

“Lack of replication for the pH treatment. Strictly speaking, substrate effect is testable, but pH effect is not formally testable because pH has only one experimental unit (header tank) per level. The pH × substrate interaction is also problematic because pH lacks replication at the level at which it was applied. Overall, shell ash was replicated among buckets and its effect can be evaluated. In contrast, pH was manipulated at the header-tank level with only one header tank per treatment, so pH effects and pH × substrate interactions should be interpreted as treatment-level patterns rather than formally replicated tests.”

Response: We redid these analyses using GLMMs to test for a header tank effect and a “header x substrate” interaction. See section 3.2 Seawater chemistry, which was completely rewritten to describe the new analyses. Linear mixed effects regression models suggest that pore water pH changed both as a function of sediment type and header tank but there were no significant interactive effects. Alkalinity was significantly affected by the header tank (p-value = 0.023), and saturation state was affected by sediment type (p-value = 0.029). We believe the revised text accurately describes the results more accurately and with greater nuance than the original version.

“2. Low replication of the shell ash treatment. Another important limitation of the experiment is that there are only 2 buckets per treatment combination. So the effective replication for testing pH and substrate effects is low. The confidence intervals and p-values should be interpreted cautiously because they are based on only 8 experimental units.”

Response: We redid the analyses and re-wrote the text of the relevant section (3.3 Shell morphology and growth) using linear mixed effects regression models. In these new models, percent new growth was statistically significant due to overlying pH (p-value = 0.042) but not presence/absence of shell hash, and the interaction between these treatments was marginally nonsignificant (p-value = 0.11). Again, we feel the revised text describes the results with the proper caveats.

“3. The statistical models used are unclear and probably unappropriated. The statistical approach does not appear to reflect the hierarchical structure of the experimental design. The authors compared several ANOVA models, including additive and interaction models with tank ID treated as a blocking factor, and selected the preferred model based on AIC. However, tank is not a blocking factor but rather the experimental unit to which treatments were applied. Individual clams within a tank are therefore not independent observations.

Given the nested structure of the data (clams within tanks), the analysis should be based on a general linear mixed-effects model (GLMM), with pH and shell-ash treatment specified as fixed effects and tank included as a random effect to account for the non-independence of observations within tanks. Model selection among different ANOVA formulations does not address this issue and may lead to pseudoreplication. Line 260: For whatever reason, the variable 'new growth' was analyzed using GLMM, which seems appropriate.”

Response: All of the older analyses have been removed and replaced with GLMMs. Section 2.9 (Statistical analyses), section 3.2 (Results: Seawater chemistry) and section 3.3 (Results: Shell morphology and growth) have been re-written to reflect the new approach. Figures 3 and 4 have also been replaced to reflect the new analyses.

“More broadly, the statistical analyses are not presented consistently across response variables, making it difficult to evaluate the strength of the evidence and compare results among parameters. In addition, the manuscript does not provide statistical analyses for the carbonate chemistry parameters measured in the overlying water, a summary table for carbonate chemistry parameters in pore water, or the full test statistics associated with the growth responses.

I recommend that the authors provide a comprehensive summary table reporting the statistical results for all measured variables. At a minimum, this table should include the model used, sources of variation, numerator and denominator degrees of freedom, effect estimates and test statistics (F-values and p-values). Such a table would greatly improve transparency, facilitate interpretation of the results, and allow readers to assess the consistency of treatment effects across response variables.”

Response: After standardizing the analyses using GLMMs, we concluded that the data would be easiest to interpret as figures as opposed to tables. Figure 3 now shows all parameters of seawater chemistry over time, and Figure 4 provides bar plots showing all measurements of clam morphology. If this Reviewer still believes a table would be useful, we would be happy to add one.

“In the same vein, the statistical treatment of the carbonate chemistry parameters could be improved. Rather than conducting separate analyses for pore water and overlying water, the authors could consider fitting a single model including water type as an additional factor. Such an approach would allow formal testing of differences between water compartments while accounting for the hierarchical structure of the experiment. Depending on the sampling design, water type could be treated as a repeated measure or as a factor nested within tank.”

Response: In the revised manuscript, we focused on pore water chemistry, since this was the subject of our hypothesis testing. We reanalyzed all aspects of pore water chemistry using a GLMM that explicitly accounts for the repeated sampling structure of the experiment. This respects the nested and repeated-measures nature of the data.

“In addition, it is unclear how temporal replication was handled in the analyses. If measurements collected at different sampling dates were treated as independent replicates, this would constitute pseudoreplication because observations through time within the same tank are not independent. Instead, sampling date should be incorporated explicitly into the model, for example as a repeated-measures factor (or fixed effect), with an appropriate covariance structure to account for temporal dependence. The authors should clarify how temporal replication was handled and revise the analyses if repeated observations were treated as independent replicates.”

Response: Our new models incorporate time into the water-chemistry models and use an appropriate structure to account for temporal dependence. Please see the revisions in section 2.9.

“Finally, I do not understand how to interpret Table 4. Also, I cannot figure out which treatments differed from which. Could you please add different letters to indicate significant differences directly in Figure 4?”

Response: Table 4 has been removed from the manuscript.

“Growth parameters and husbandry conditions are difficult to appreciate. As a bivalve biologist, I was struck by the small increments in shell size observed after 90 days at a temperature of 15°C. Is this normal for this species? I would appreciate a discussion about this specific point. How does this growth compare to what happens under natural field conditions?”

Response: There is not much work on growth rates in *Leukoma staminea*. Our previous work suggests they reach ~7-15 mm after 200 days (Kempf et al. 2023), and specimens from British Columbia reach 22 to 35 mm long between 1.5 and 2.5 years old: <https://www.dfo-mpo.gc.ca/species-especes/profiles-profil/littleneck-clam-palourde-pac-eng.html>. So these do not appear to be fast-growing animals. The juveniles in this study were already 200 days old when the 90-day experiment began, so the measured growth reflects post-settlement laboratory growth under controlled conditions rather than early larval development. See lines 393-395 in the revised text.

“Discussion needs further development. The discussion is far too short. I would very much appreciate some thoughts on the potential of using shell ash to cope with ocean acidification. Do you think this is feasible? Is it scalable? Would it be acceptable to growers and fishermen? Furthermore, given the limitations of the experimental design, I would be grateful for a statement on this in the discussion.”

Response: We have added three paragraphs to the Discussion to better consider these issues. In short, we conclude that shell hash can buffer acidified sediments and improve conditions for bivalves, but the biological benefits are inconsistent and depend strongly on factors such as species, life stage, predators, vegetation, and local environmental conditions. Overall, shell hash appears most promising as a short-term, site-specific habitat-engineering tool rather than as a broadly effective mitigation strategy.

“Please respect the following variable order in the manuscript text, figures and tables: pH, TA and omega”

Response: Done

“Figure 3 plots pH_{tot} in pore waters as a function of time for each treatment. This is the only variable presented with such detail. Why is that? Please consider doing the same for other carbonate chemistry parameters, at least in a supplementary file. If not, please provide a justification. It is likely that the TA will

increase over time in the shell ash treatment, particularly under OA conditions. In any case, time should be investigated and considered a repeated measure in the statistical model (see point 3)."

Response: The new figure includes all carbonate chemistry parameters.

"Add some description of the shell ashes composition, grain size, ..."

Response: Please see lines 154-160

"F-values in the text are not well reported. The author reported as F(numerator df) while it should be F(numerator df, denominator df)."

Response: These have been removed from the text.