Reviewer 1

The authors present the findings from an ocean alkalinity experiment (OAE) addressing the response of larval herring. Two separate experiments were carried out: small-scale experiment on individual larval herring response, and a large mesocosm experiment designed to examine community level interactions in the context of OAE. The authors report non-significant effects of OAE on herring larvae in both experiments. This manuscript provides some baseline evidence addressing the importance of identifying potential negative impacts of OAE. Several aspects of the data analysis and experimental design need additional details and clarification before publication.

We are grateful to both reviewers for their very thoughtful and thorough feedback. Its implementation improved the quality and clarity of our manuscript significantly.

While addressing the reviewers' comments, we became aware of a shortcoming in the laboratory experiment. The pCO_2 concentration in the 'control' fish tanks was above what one may consider 'ambient'. It appears that respiration in the tanks was not entirely compensated for by the aeration system, leading to an accumulation of CO_2 and thus acidification. With that we lose our 'control' treatment for the test of OAE. The laboratory results cannot be contrasted to the ones from the mesocosm. We removed the laboratory part from the manuscript.

This re-orientation did not alter the principle take-home message of the study. Changes were required at a more detailed level, however. Specifically, the 'mineral' treatment of our mesocosm experiment is now properly integrated. Before, it had been restricted to the method section to allow for a straight forward combination of mesocosm and laboratory experiments. Judging from the reviewers' comments, this neglect of the mineral treatment had been confusing. So, in a sense, the now uncompromised focus on the mesocosm improves clarity. In the revised manuscript, the structure and data presentation match the experimental design exactly. The updated title and figures are provided as supplement at the end of this document.

General comments

1) The introduction appears to focus more on pH and pH variability rather than OAE. As pH was not held constant during these experiments, the authors should shift the focus of the introduction to alkalinity specific effects. This could take the shape of effects of added Ca with OAE and osmoregulation. The introduction also lacks details about the concerns of OAE and the reactions that occur. More attention needs to be given to this to build context.

The increase of alkalinity itself is not affecting biology. Also, changes in osmolarity are minor compared to the osmolarity of seawater and have thus been consider negligible as stressor (stated in line 168). Instead, biology is expected to be most strongly driven by shifts in associated variables (e.g. pH, pCO_2, ΩCa). These are inseparable from alkalinity and thus an intrinsic element of OAE. In our opinion, an understanding of their effects is what is required to evaluate environmental safety of OAE. We give the associated increase in pH particular attention as it may matter most for fish physiology. Please note as well that the non-CO_2 equilibrated OAE approach tested here entails particularly strong shifts in carbonate chemistry (pH, pCO_2, ΩCa), which makes their study especially relevant.

We state the above more clearly in the revised manuscript:

"Biological processes are not affected by alkalinity itself but through the associated changes in various ions and molecules (Bach et al., 2019)."

The general explanations around OAE including chemical reactions, associated changes in carbonate chemistry and the various implementation scenarios of OAE are already presented several times in great detail in the

existing literature (e.g. cited here: Renforth and Henderson 2017, Bach et al. 2019). In contrast, information (theoretical and empirical) on the potential consequences for fish is not yet available. To increase the value of our manuscript (via complementarity), we prefer keeping the general introduction on OAE short so that more focus can be placed on introducing the questions that matter more specifically for fish.

2) What is the reason for choosing such a moderate delta TA (600 umol kg-1). Predictions of real applicable scenarios estimate higher TA (See Renforth and Henderson 2017).

In the last paragraph of the introduction (lines 79-82), we classify our maximum delta TA of 600 as relatively high and provide references for it: Bach et al., 2019; Hartmann et al., 2023; Renforth and Henderson, 2017. What matters here is whether an equilibrated (CO_2 added together with alkalinity) or non-equilibrated (CO_2 sequestration happens afterwards) OAE application is used. We here test the non-equilibrated application, which is the basic approach to OAE. This is stated in the aims paragraph of the introduction (line 79) and the treatment section of the methods (line 118) and context is given in the introduction (lines 42-45) and discussion (lines 324-327).
The non-equilibrated application involves much more drastic increases in calcium carbonate oversaturation at any given delta TA than the equilibrated one. In other words, abiotic precipitation is occurring sooner. Our choice of maximum delta TA was meant to (just) avoid precipitation. The difference between these two approaches and the risk of abiotic precipitation is a central topic in the references cited, including Renforth and Henderson 2017.

Our original manuscript obviously lacked clarity here. The 'OAE application' section in the methods of the revised manuscript now includes a thorough explanation.

"ΔTA was applied non-CO_2 equilibrated (Bach et al., 2019). This economically more feasible approach, only adds alkalinity and CO_2 sequestration occurs afterwards through natural equilibration with the atmosphere. In the alternative pre-CO_2 equilibrated approach, the to-be-sequestered CO_2 would instead be added together with the alkalinity leading to milder changes in seawater chemistry. We restricted ΔTA to 600 µmol kg-1 to avoid abiotic precipitation of calcium carbonate, which would signify a loss in alkalinity and thus a nonsensical OAE scenario (Hartmann et al., 2023)."

3) Since pH was not held constant, this should be considered as a random effect in the mixed model. Despite the finding of nonsignificant effects with the current model, the added random effect will shift the distribution of your df giving you a more accurate representation of detla_TA as a predictor variable.

Our fish responses represent a temporal integration of abiotic and biotic conditions. This is especially the case for variables like growth and survival but also physiology and behavior to some extent. Including current pH as covariate would not reflect this lag in response. The lag is unknown and will differ between variables. There could even be effects in the history of the larvae that are irreversible. Then, strictly speaking one cannot stop at pH but needs to consider other carbonate chemistry variables that are modified by delta TA like pCO2 and ΩCa. Our experimental design is not able/meant to disentangle their effects. Instead, our non-CO_2 equilibrated deltaTA levels should be seen as 'OAE scenarios' that entail a certain carbonate chemistry (pH, pCO_2, ΩCa etc.).

To note, temporal variability in carbonate chemistry is less of a topic in the revised manuscript that is based on the mesocosm only. These variables were relatively stable in the mesocosm. The revised manuscript shows the temporal development of not only of pH but also pCO_2.

4) Why was it chosen to have no replicates for the mesocosm experiment? I understand seeing no mineral effect in the lab experiment as potential justification but it was never mentioned that Si was used during the lab experiments. How do you

know Si didn't carry a random effect. This needs to be explained and justified.

The revised manuscript is less confusing in this respect as it only includes the mesocosm design without compromises. The mineral treatment is now fully integrated in all data analyses and their interpretation. In the 'OAE application' section, we also added an explanation regarding the choice of a gradient design for ΔTA.

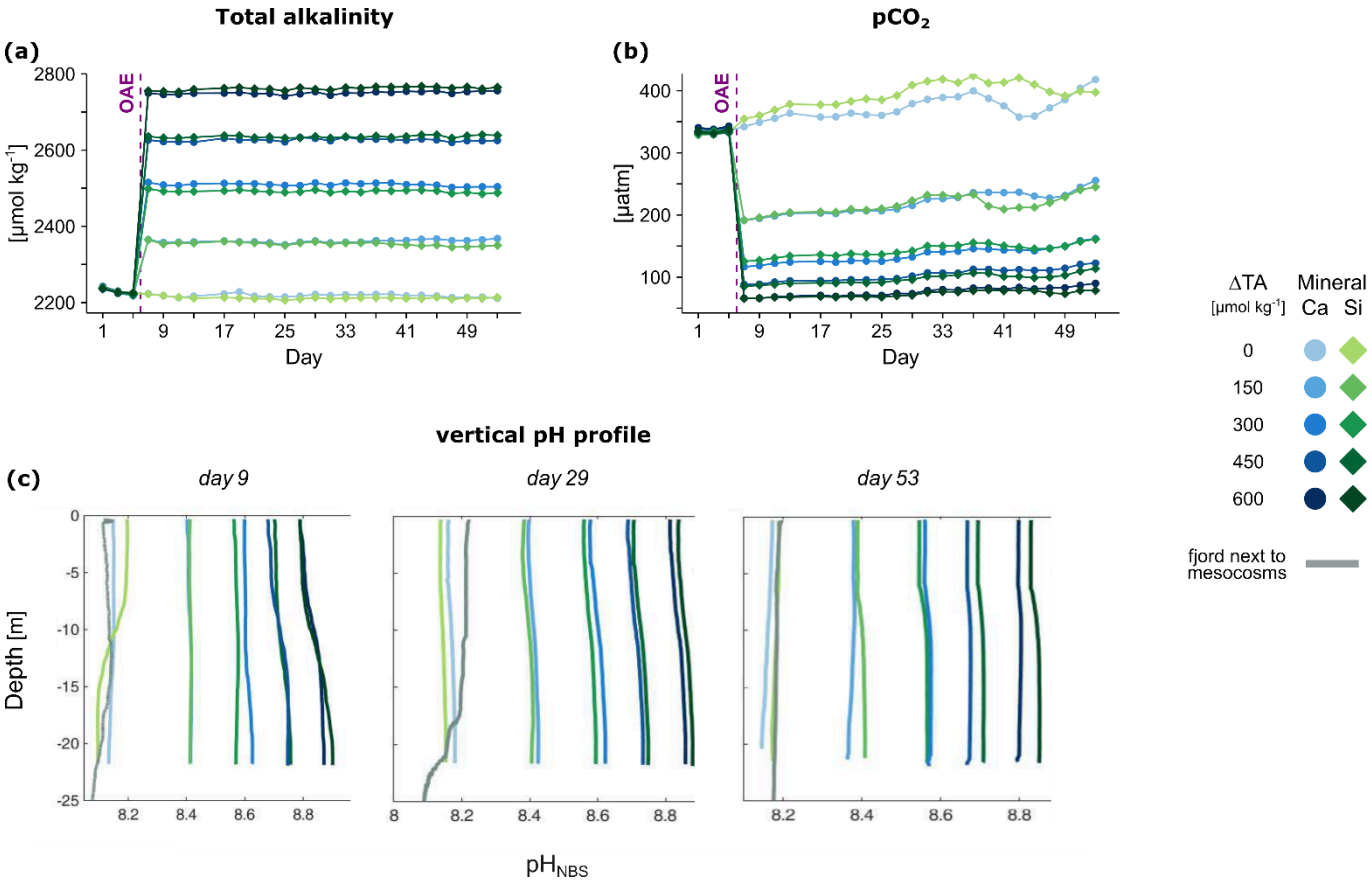
"The gradient design with non-replicated treatment levels (Riebesell et al., 2023) was preferred for ΔTA to allow for a more informative study of biogeochemical processes that were also part of the multidisciplinary mesocosm project (e.g. Ferderer et al., 2024). For analysis, ΔTA could then be tested as continuous explanatory variable in the sense of linear regression."

5) There is no timeseries of delta_TA just pH. A delta TA timeseries needs to be shown. If this information is not available, then report and justify why not. Where was pH measured in the mesocosms, surface, depth? If stratification was strong, there could have been a pH gradient as well, particularly just from the temperature effect from 8 to 15 C. This needs to be address? Again, pH is a covariate for the mesocosm experiment. Discuss this.

1) Yes, we agree with the need for a full presentation of the carbonate chemistry data. The revised manuscript includes a supplementary figure with the temporal development of TA and pCO_2 as well as a vertical pH profile. pH of time is in the main text.

TA is not shown in the main text as it corresponds to the intended target values already provided via the experimental design and figure legends. Unless there is precipitation, TA is stable during OAE.

The variability in pH across depth was small compared to the shift in baseline caused by the TA manipulation itself. An elaborate discussion is not needed. We added a corresponding sentence in the main text in the "OAE application" section:
"These OAE-induced shifts in carbonate chemistry were present across depth (as exemplified via pH, Fig. S1c)."



"Figure S1: Further assessment of carbonate chemistry. Development of **a)** total alkalinity (TA) and **b)** pCO2 in each mesocosm unit. **c)** Depth-dependent variability in pH. Sampling days at the beginning, middle and end of the treatment period serve as examples. These pH measures were taken in situ via CTD with a potentiometric pH sensor (NBS scale) and are hence slightly higher than the spectrometric measures (total scale) shown in figure 1c."

2) In the revised manuscript, we provide a more detailed explanation about how water was sampled:

"Carbonate chemistry and inorganic nutrients were monitored in two-day intervals based on depth-integrated water samples. For this, samplers equipped with a pressure-controlled motor (5 L, Hydro-Bios, Kiel, Germany) where lowered from the surface to the bottom of the mesocosms to collect water evenly across the water column. The resulting samples represented mesocosm averages."

6) For the NMDS plots, it would be better to examine larval herring survival or biomass to other assemblages, chl a measurement, temp and OAE rather than changes to whole assemblages of other fish. If you also want to examine the effect of different fishes, then examining the different environmental variables would be useful. Could be useful to also show correlations with other variables to explain the lack of effect of OAE.

1) We decided to not analyse the final fish assemblage from the perspective of the herring as they were not dominant amongst the fishes (see Fig. S3). Instead we took a more neutral stand considering all fish groups equally. The question we address with this multivariate analysis – is there evidence for an OAE-driven shift in the composition of the fish assemblage? - is basic but valid. Under different circumstances (high herring survival dominating the fish assemblage), we would agree with the analysis strategy recommended by the reviewer.

2) Our study is based on an experiment with only 10 independent measurements (via 10 mesocosms). Unfortunately, that means we are limited in the number of predictors that can/should be included in the model due to overfitting. This is true for both our multivariate and univariate analyses. We hence restrict ourselves to the predictors of our experimental design, as only these test cause and effect. Following *deltaTA*, the next predictor is *Mineral* and the interaction *deltaTA x Mineral* (see Table S1). With that we have come close to the limit of our data structure. Again, under different circumstances (more independent measurements), we would agree with the reviewer in also considering (correlation) other abiotic and biotic predictors.

Line comments:

39: List some biological processes specifically

This first sentence is the topic sentence of the paragraph. The arguments come after. Most of our paragraphs follow this assert-justify style of writing.

In this case, the next sentence (lines 39-43) lists the biological processes:

"Added alkalinity reduces CO_2 as carbon source for primary producers (Hansen, 2002), rises calcium carbonate saturation which facilitates calcification (Renforth and Henderson, 2017) and increases pH concerning the acid-base balance of organisms (Tresguerres et al., 2020; Pörtner, 2008).

45: provide reference for gigaton estimate and scale of plume/ deployment mass of material

In the revised manuscript, we reference Renforth and Henderson (2017) here. This general review on OAE covers precisely these points. Further, aspects are covered by He and Tyka (2023) and Bach et al. (2019) that we cite already throughout the paragraph.

46: This line doesn't make much sense. ''because of this coastal systems are relevant". Provide more description/context.

The context is given in the subsequent sentence.

“They are not only most attractive economically for OAE deployment given the proximity to mineral and energy sources (He and Tyka, 2023) but also hotspots of biodiversity and ecosystem services.”

56: What are these thresholds. Provide more discussion here. Where these things even specifically addressed in the study?

We mention “energetic costs” and “thresholds” (that we indeed don’t study specifically), as these are key processes that can cause altered growth and survival (which we study) in the context of stressors in general. Specifics about thresholds for fish under OAE and the underlying mechanisms are not yet described in the literature, given the novelty of the field. In the revised manuscript, we adopted a more neutral formulation of this content:

“There could be additional energetic costs for acid-base regulation that channel resources away from growth and reproduction or pH thresholds beyond which physiological functions fail and threaten population viability.”

78: indirect effects can be explained better.

“Indirect effects” were introduced in the preceding paragraph

lines 70-72: “At the community-level, larvae are tightly controlled by resources, competition and predation (Houde, 2008). OAE could change these food web interactions, for example via species-specific pH sensitivities, expanding calcifiers or CO2 limited primary production, giving rise to indirect effects (Ockendon et al., 2014; Goldenberg et al., 2018).”

and explained again graphically in the same paragraph via figure 1b (line 75).

105: Is this 1800 larvae in total, so evenly distributed across all treatments?

In the revised manuscript, this sentence now only relates to the larvae rearing for the mesocosm introduction, given that the laboratory experiment was removed. We made sure to be clearer this time:

“After hatching 17 days later (24th April), a total of 2400 larvae were distributed amongst two ~500 L rearing tanks.”

106: Was temperature maintained and continuously monitored in the aquaria?

We clarified this in the revised manuscript:

“Temperature was continuously adjusted to match that of the outside fjord and averaged 6.5 ± 0.8 °C.”

115: Explain better. Was this a column filled with soda lime ?

Yes, it was. But this information is not relevant any more now that the laboratory experiment is removed.

116: how was PCO2 measured?

In the revised manuscript, we now include the method for the pCO_2 calculation. To note, this relates to the mesocosm only.

“Carbonate chemistry was assessed following Schulz et al. (2023). TA was measured via titration (Metrohm 862 Compact Titrosampler with Aquatrode Plus with PT1000) calibrated against certified reference material (CRM batch 193) supplied by Prof. Andrew Dickson’s laboratory and pH spectrophotometrically (Dickson et al., 2007). With the temperature and salinity provided by the CTD casts, in situ pH and pCO_2 could then be calculated using CO2SYS for Excel with constants from Luecker et al. 2000 and Dickson 1990 (Pierrot et al., 2021).”

118: Show this TA time series and variation across reps.

The full TA development is now shown in supplement for the mesocosm experiment.

123: How was this pH probe calibrated? Tris, NBS buCers? What pH scale (report this throughout)

Thanks for pointing out the lack of pH scale. In the laboratory, pH had been measured potentiometrically on the NBS scale and was then converted to the total scale. In any case, this is not relevant any more now that the laboratory experiment is removed. For the mesocosm experiment, pH scales are now reported throughout the manuscript.

133: So a subsample of 5 larval fish were measured per aqauria. Out of how many total in each aquaria?

Not relevant any more due to removal of laboratory experiment.

152: Show depth gradient in tanks form the CTD data. Where was pH measured. At what depth. Where was TA measured? What depth? Do you have any depth resolution?

The revised manuscript includes a figure with the pH depth gradient measured via CTD. There is no depth data available for TA and also no reason to expect major variability here given the pH profiles. While the pCO_2 depth gradient could in theory be calculated using this information, we believe this to exceed the scope of our manuscript on fish.

We now also include a better description of how water was sampled:

"Carbonate chemistry and inorganic nutrients were monitored in two-day intervals based on depth-integrated water samples. For this, samplers equipped with a pressure-controlled motor (5 L, Hydro-Bios, Kiel, Germany) where lowered from the surface to the bottom of the mesocosms to collect water evenly across the water column. The resulting samples represented mesocosm averages."

170: Why then were OAE duplicates two different minerals. If Si mineral addition creates reduced nutrient scenarios then having one replicate at one level of delta TA gives you no power. This needs more discussion.

ΔTA is manipulated as continuous explanatory variable and can so be tested in the sense of regression statistics. Here, the 'replication' needed to separate effect from noise is coming from the 5 independent measures along the gradient. Replication of a given level is not required and actually not wanted. Such a design with a continuous explanatory variable is common in ecology (e.g. Quinn and Keough 2002 Experimental design and data analysis for biologists). Its popularity equals that of a factorial design. The statistical calculations underlying these two approaches are the same. We are hence not including a full explanation.

Still, it seems that some more background is needed, which we provide in the 'OAE application' section:
"The gradient design with non-replicated treatment levels (Riebesell et al., 2023) was preferred for ΔTA to allow for a more informative study of biogeochemical processes that were also part of the multidisciplinary mesocosm project (e.g. Ferderer et al., 2024). For analysis, ΔTA could then be tested as continuous explanatory variable in the sense of linear regression."
The reference 'Riebesell et al' includes a section dedicated to design choices in mesocosm studies. It discusses the 'gradient' approach and associated statistics.

The 'Data analysis' section describes the model structure: "… linear models were employed with ΔTA (continuous), mineral (categorical) and their interaction (ΔTA × mineral) as explanatory variables (type III test)."

174: Explain the effect of NaSiO as means to prevent secondary precip. Discuss the reactions.

There may have been a misunderstanding, sorry. Na_2SiO_3 provided the silicate for the Si-based mineral treatment. We restructured the sentence:

"Silicate was added in equal amounts of 75 µmol L^{-1} to all five Si-based mesocosms using Na_2SiO_3. This allowed to separate the effects of alkalinity and silicate and prevented mineral precipitation (Ferderer et al., 2024)."

The provided reference is from the same experiment and dedicated to the topic of silicification. It has more information on the risk of secondary Si precipitation.

200: See general comment above about LMM

The laboratory analysis part was removed from the revised manuscript.

216: this should be in the discussion.

The laboratory results were removed from the revised manuscript.

220: If you want to shift the focus to post-exposure effects and discussion then their needs to be some alignment in the discussion about this, despite finding no effect. There are several OA experiments looking at post-exposure effects, these can be used as a contrast.

This is not relevant any more due to the removal of the laboratory experiment. In the mesocosm experiment, the stressor levels are constant throughout the experiment.

227: What is standard herring survival rates for natural communities?

Not relevant any more, as part of laboratory results.

241: How were carcasses differentiated between species?

The revised manuscript includes this information in the method section:

"For this, the sediment trap was sampled in two-day intervals via a tube connected to the surface (Fig. 1a) and immediately screened for dead fish. Carcasses were assigned to either herring, codfishes or flatfishes based on their distinct body shape. This method had proven successful in previous campaigns, especially in colder climates where fish carcasses disintegrate slowly (Spisla et al., 2022).

303: Describe the food-web mediated impacts of OAE."

The entire paragraph is about the food web-mediated impacts. After the paragraph's topic sentence (line 303), we first remind about some of the potential effects proposed by the literature (lines 303-306) and then follow with our mesocosm results (lines 306-313).

308: Given you had no real replication of a specific mineral at high TA, an NMDS associating fish biomass with other drivers specific to each mesocosm could help explain this.

Please see the discussion above regarding the non-replicated design and the testing of other drivers via correlation.

315: provide more discussion of other metrics used to determine herring tolerance and survival. More discussion needs to be given about why OAE didn't effect them. What are their osmoregulation strategies, what are the potential, for carry-over effects, etc..

The thorough testing of direct effects of OAE on herring from the laboratory, including metabolic rate and behavior, is not included in the revised manuscript. The focus has shifted to the community level. While these traits still matter for the growth and survival in the mesocosms, they are now not specifically tested any more. We believe that our more general introduction (lines 49-66) and discussion (lines 283-302) on physiology at the organism-level is now appropriate.

The past development of research on other stressors (e.g. ocean acidification) shows a strong preference for laboratory studies on direct effects on physiology. Community-level studies are instead rare because of their large cost. We hope that future OAE research will test and discuss physiological responses in detail and complement our work.

Figure 4 and 5 need a legend for the color.

Color legends now included.

Appendix

title: Viability of fish larvae under ocean alkalinity enhancement in coastal plankton communities

all figures of the revised manuscript:

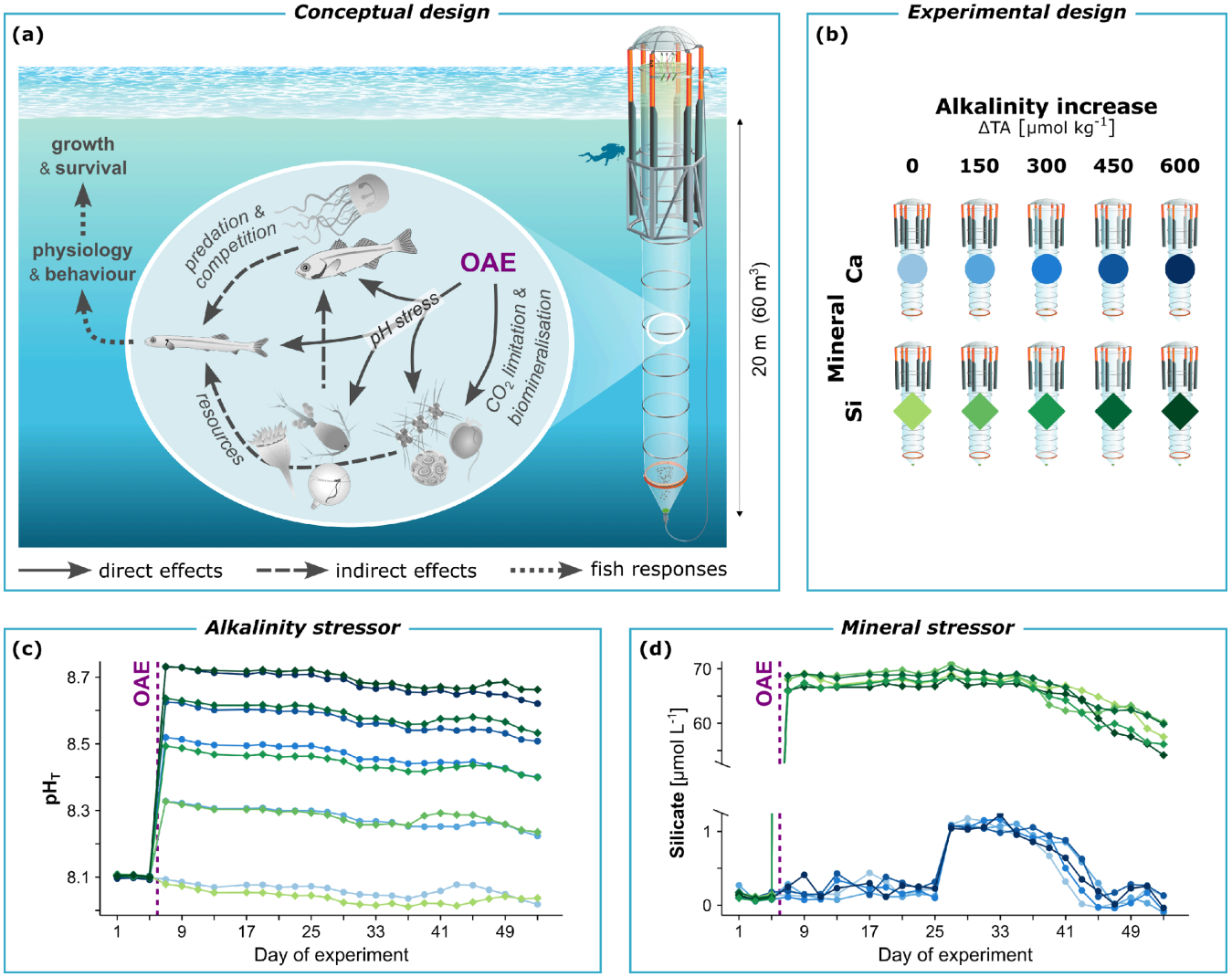


Figure 1: Conceptual and methodological framework of our mesocosm study on non-CO_2 equilibrated OAE. **a)** Potential pathways of change in fish in natural plankton communities. **b)** Water chemistry manipulations to simulate different scenarios of OAE. Using 10 mesocosm units, we tested increases in total alkalinity (ΔTA) under calcium-based (Ca) or silicate-based (Si) mineral addition. **c)** OAE-induced shifts in pH and **d)** silicate availability, as measured in each mesocosm throughout the experimental period. Mesocosm symbol from Rita Erven, GEOMAR, and organism symbols partly from Integration and Application Network, University of Maryland Center for Environmental Science (ian.umces.edu/symbols/).

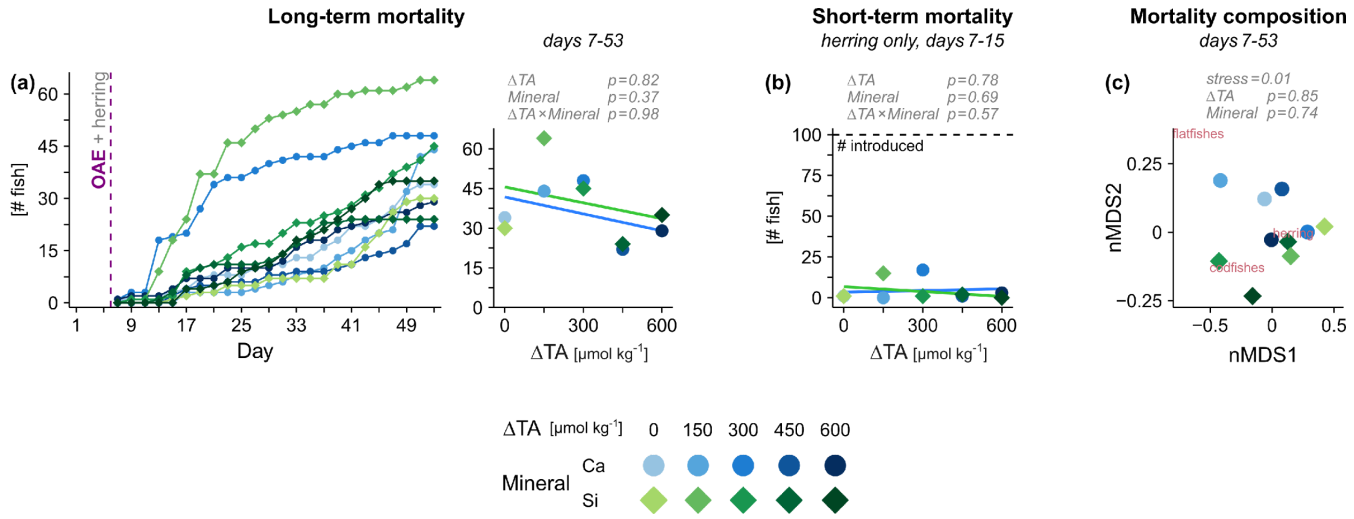


Figure 2: Fish mortality under OAE monitored via the sediment trap. **a)** Cumulative mortality over time across all species and **b)** immediately following the OAE perturbation for herring. **c)** Differences in taxonomic composition between mesocosms via non-metric multidimensional scaling (nMDS). Statistical tests in grey (details in Table S1).

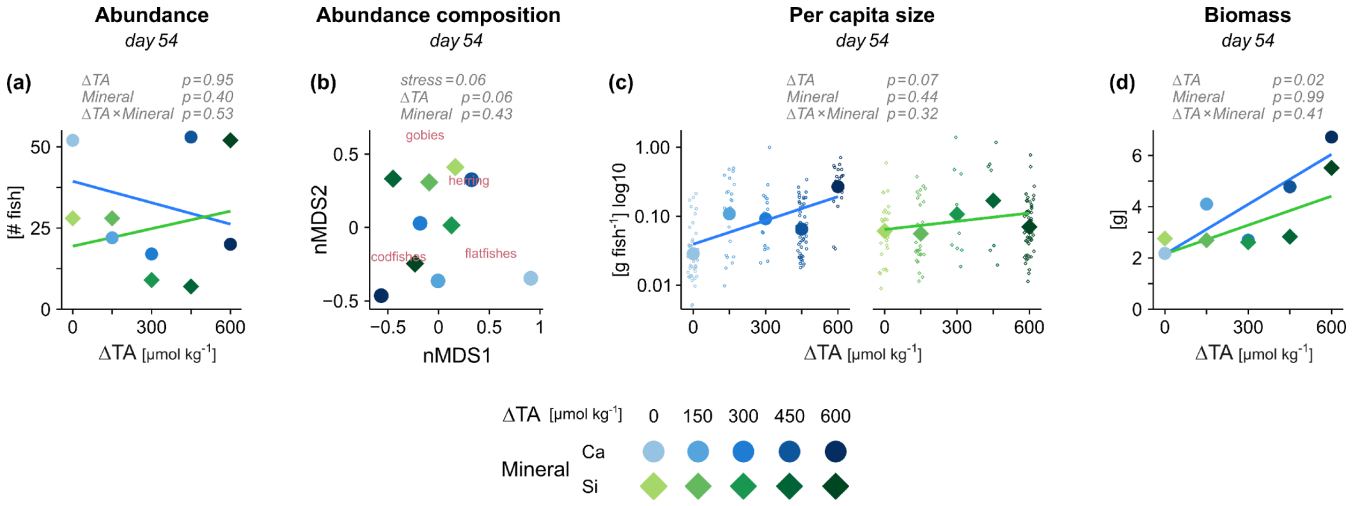


Figure 3: Fish growth and survival under OAE, assessed via the assemblage of live fish at the end of the experiment. **a)** Abundance, **c)** individual size and **d)** total biomass across all fish taxa. **c)** Differences in taxonomic composition between mesocosms via non-metric multidimensional scaling (nMDS). Larger points represent mesocosms and smaller points in **c** single individuals. Statistical tests in grey (details in Table S1).

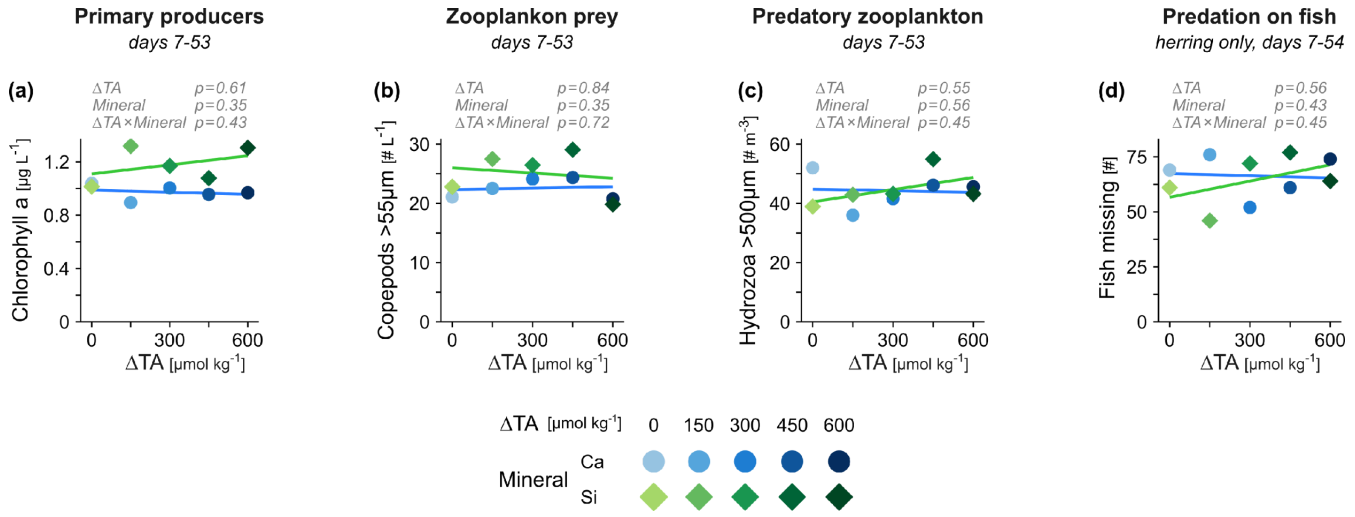


Figure 4: Potential sources of indirect effects of OAE on fish mediated via species interactions. Abundance of other functional groups including (a) primary producers, (b) invertebrate grazers and (c) invertebrate predators. d) Predation on herring estimated via missing individuals. Averages across the treatment period are tested (in grey, details in Table S2).