First of all, we would like to thank the referees T. Tyrrell and A. Dickson for their detailed and constructive comments on our manuscript. Changes in the manuscript are highlighted in blue.

Specific comments by Toby Tyrrell

Pg 1782, line 25: We agree and have added the information given by Canadell et al. (2007).

Pg 1783, line 13: We added the important remark that “in many experiments, the
carbonate chemistry is not at all controlled after initial adjustment”.

Pg 1783, line 18: We added the point that the combination of the parameters TA and DIC “had also been thought to lead to the most accurate calculations of CO2 concentrations and carbonate saturation states (Riebesell et al. 2010)”.

Pg 1783, line 27: We have based the statement that “the quality of carbonate chemistry measurements at high pCO2 is often not known” on the fact that CRMs are only available for ambient pCO2 rather than on the results of our study. We have specified this in the revised manuscript.

Pg 1784 results section: At the time of writing the current version of the manuscript, there were no datasets found in the EPOCA database showing consistent carbonate chemistry (i.e. calculated from three parameters) with data originating from OA laboratories. We have stated this more clearly in the revised manuscript.

Pg 1784, line 26: We agree and have specified this in the revised version. In fact, we were referring to lacking information on nutrient concentrations. In case of the study by Schneider & Erez (2008), nutrient concentrations were taken from Makey et al. (2009, 0.1 μmol kg-1 phosphate and 2 μmol kg-1 silicate). In case of the dataset from Thomsen et al. (2010), concentration of 2 μmol kg-1 phosphate and 3 μmol kg-1 silicate were assumed (Gerlach, 1990). In case of the dataset of Iglesias-Rodríguez et al. (2008), typical concentrations in (non-diatom) phytoplankton cultures were assumed (8 μmol kg-1 phosphate and 6 μmol kg-1 silicate). In this range of nutrient concentrations, changes in carbonate chemistry due to different nutrient concentrations are negligible compared to the discrepancies described.

Pg 1785 lines 6-9: Based on the information on this dataset, outgassing cannot be excluded as an explanation for the discrepancies between expected and measured pCO2. Because the discrepancies show exactly the same pattern as the other datasets, we think this explanation is rather unlikely. We have, however, mentioned this possibility in the revised manuscript.
Pg 1785 line 14: We have restructured the paragraph and added the requested information on how well the different pCO2 values compare.

Pg 1785, line 21-23: All studies mentioned have compared measurements and calculations from a range of pCO2 levels (approx. 200-1800 µatm), falling well within the range of those described in this study. We have specified this in the revised manuscript.

Pg 1876 lines 4-5: These values are in fact based on repeated CRM measurements. We have stated this in the revised manuscript.

Pg 1786 lines 17-18: Unfortunately, we have not measured pH before the AT manipulation of seawater carbonate chemistry to test this idea. The differences between initial and expected final AT (i.e. after addition of HCl/NaOH), however, met the expectations within 10 ±2 µmol kg⁻¹ (uncertainty deriving from two sets of AT measurements and pipetting errors). Therefore, the discrepancies cannot be explained by uncertainties in AT measurements alone.

Pg 1787 lines 18-19: We thank the reviewer for highlighting this point and have added that care must be taken “when reporting threshold levels of pCO2 harmful to an organism”.

Pg 1788 line 5: We agree that a realistic estimation of uncertainties involved in the measurements of the different carbonate chemistry parameters is not straightforward, especially when systematic errors occur. From our point of view, it would be useful to establish a certain routine of estimating (i.e. by means of regular CRM measurements), stating and propagating uncertainties, irrespective of their origin. Furthermore, together with Prof. A. Dickson, we are currently planning an inter-laboratory comparison of carbonate chemistry measurements (especially at high pCO2 levels). So far, about 70 laboratories (involving both marine chemists and OA researchers) have stated their interest, hopefully helping to reveal whether the phenomenon described here is indeed widespread. We have mentioned the need for such an exercise in the revised manuscript.
Fig. 1: We thank the reviewer for pointing out this inconsistency in the use of abbreviations and have changed the axis labels and stated more clearly that the results are shown for 15°C for all datasets.

SI Pg line 22: Wording changed to “The CO2 concentration was regularly measured with a non-dispersive infrared analyzer system“.

SI section 4: We agree and have changed the text accordingly. Repeatability and reproducibility describe different conditions of measurements on which estimations of uncertainty can be based (cf. Taylor & Kuyatt: Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results, NIST Technical Note 1297, 1994).

SI pg 3 line 22: We have corrected the typo and added the reference to the revised SI.

Specific comments by Andrew Dickson

(1) We fully agree that “it should be common practise to provide and defend an estimate of uncertainty” and have explicitly stated this in the revised manuscript.

(2) We consent with Prof. Dickson that agreeing on two parameters will not allow progress in terms of the “true” carbonate chemistry. In order to achieve better comparability between different studies, however, it might be useful to agree on a certain pair of input parameters. This will not, as Prof. Dickson pointed out, guarantee complete comparability but nevertheless improve it. With respect to the choice of a certain pair of parameters, we favour pH and AT, even though DOM and other acid-base systems can significantly contribute to AT. This, however, is mainly a problem in very dense cultures, which should be avoided in OA studies anyway. AT is the only parameter that is not changed by outgassing of CO2 during sampling and storage, which in our opinion, can be an even bigger issue in OA research.

Interactive comment on Biogeosciences Discuss., 9, 1781, 2012.