

**Anonymous Referee #1:**

We thank Referee #1 for this short positive review. For the final manuscript we will change the text on page 6630 according to the suggestion.

**Anonymous Referee #2:**

- The number of participating laboratories was 13. The corresponding information will be made consistent. About half of these laboratories provided measurements for the full set of flasks obtained over time, including the sample flask with ambient air. Hence, the lower number of 6 laboratories, from which the results have been compiled and presented.

- The missing reference on pages 6628 and 6633 will be added.

- The suggestions for page 6628 L6 and L10 will be taken into account.

- The footnote on P6629 L28 will be changed accordingly

- P6630 L7: The calcites Mar-J1 and OMC-J1 are both similar to NBS 19. The materials have been described in Ghosh et al, 2005. A note will be added.

- P6630 L12. The issue of alteration of CO<sub>2</sub> isotopes in large flasks has been discussed extensively before (also in Ghosh et al, 2005). However, the sentence referred to in the comment related to the stability of the scale anchor arising from the fact that the stored material is calcite, not a gas.

- P6630 L22. Suggestion will be followed.

P6632 L26. The sentence refers to the common observation that d18O uncertainties of CO<sub>2</sub> isotopic measurements usually are larger, which is attributable to several reasons including those explicitly mentioned. The fact that the 18O abundance is smaller than that of 13C is one not mentioned, the tendency to exchange oxygen with (wet) surfaces another; the list could become very long. For the N<sub>2</sub>O correction, this also applies from the fact that the absolute correction is larger for 18O than for 13C. This has been discussed in the cited literature. A corresponding expansion of arguments has been added.

**Anonymous Referee #3:**

Referee #3 expects a lot of changes from us, which would result in a much longer and altered paper. In fact, the paper would have to go through another review, because it would be quite different. While we share that a number of specific issues have not been discussed in this manuscript, we have published about most of the points raised rather extensively and discussed the issues in the papers which have been cited throughout our contribution. We did not want to repeat ourselves by discussing everything from scratch. Rather, the new results of the JRAS comparison is what we focused on and we believe that the relevant data are properly represented and discussed and the conclusions are clear and concise. We will, however, follow most of the more specific suggestions raised and implement the changes in the final submission:

*Q: P6628L11: The sentence is not clear to me. Did you mean that mixing reference CO<sub>2</sub> in air is unique?*

*A: We don't think the process as such is unique. However, the provision of CO<sub>2</sub> generated from calcites and mixed into CO<sub>2</sub>-free air as reference material is unique, as far as we are aware of. The text has been altered for clarification.*

*Q: P6628 L15: Please spell out the "IMECC" and any acronyms at the first time.*

ok

*Q: P6629 L2-4: Please be consistent with the names of the participated labs/organization, i.e., "CAR/CSIRO" and INSTAAR/NOAA need to be spelled out. The Center for Atmospheric and Oceanic Studies should be replaced with Tohoku University (TU), which has been conventionally used for Dr. Nakazawa's*

group (Allison et al., 2003).

A: Changes have been applied

Q: P6629 L7- P6630 L2: Based on the statements in the CLASSIC report, the authors introduced that the root of the problem, i.e., the discrepancies for both pure CO<sub>2</sub> and CO<sub>2</sub> in air between labs, is scale definition. At the same time, the scale contraction caused by cross-contamination, inconsistencies in <sup>17</sup>O correction and the algorithms for N<sub>2</sub>O correction are listed as the contributed factors by the authors. It is expected that a thorough discussion to advance our understanding for those factors will be provided in the following sections to narrow down the dominant factor. But this was not done... The comparison results with different <sup>17</sup>O corrections and N<sub>2</sub>O algorithms applying to the JRAS measurements should be provided in a table (in the following sections) to infer the dominant factor causing these discrepancies.

A: These questions have been extensively discussed and answered to a large extent in the preceding and cited literature (e.g. Wendeberg et al, 2011; Brand et.al, 2009). The corresponding details were not part of the JRAS intercomparison. Rather, the aim was to relate established, yet different scales to JRAS 06 and find correction parameters. From the data it is clear that scale compression is still a dominant factor.

Q: P6630 L24 - P6631 L2: Based on the content, JRAS-06 is the local scale of MPIBGC Isolab, which is firmly anchored at VPDB scale and continuously maintained. Please describe/define the JRAS-06 scale (are the two 5L flasks made from MAR-J1 and OMC-J1 in 2006?) and how the scale is maintained over time? It would be more convincing to show the data for scale maintenance in a table from 2006-2010.

The MPI-BGC data in figures 2 and 3 provide this information. The MPI-BGC lab participated in the exercise as a routine analytical lab, just as the other participants.

Q: P6631 L14-15: Does it mean that a JRAS set consists of three flasks after the spring of 2010? It is not clear if the three-flask sets of JRAS have been analyzed and evaluated by all the participating labs.

The information why only 6 labs have been included has been given. The JRAS set is still based on the 2-flask approach with the CO<sub>2</sub> from calcites, but the air flask is included occasionally in order to detect systematic scale compression. The air flask cannot be part of the primary calibration because we were not able to identify a calcite material with the right properties, including an isotopic composition close to air-CO<sub>2</sub>.

Q: P6631 L20-21: Based on the last paragraph in section 2, the local scale at MPIBGC is JRAS-06 scale. Is the evaluation of JRAS-06 (shown on Fig.2) by the MPI-BGC local scale independent? It seems that both axes are at the same scale (i.e., JRAS-06). A linear regression has been applied to the three data points shown on Fig. 2. Could you please provide the information in a table, including all the measurements for each of the three points (the MAR-J1, OMCJ1 and the dry ambient air) and the date for individual measurements?

The reply given for the previous question also applies here.

*Q: P6632 L19- L26: Again, it is not clear if the three-flask sets have been measured by all the participated labs. If yes, please present the data ( $\delta^{13}\text{C}$  and  $\delta^{18}\text{O}$ ) which was used for the regressions (in Tables 2 & 3 and Figs.4 and 5) for each participated labs in another table. The date for each analysis should be also included for showing the evaluation of the comparison over time instead of one time exercise.*

We believe that all necessary information is given, including the linear regression analysis. The individual measurements do not add to the results.

*Q: P6632 L23: Please change the  $\text{VPDB}_{\text{gas}}$  to  $\text{VPDB-CO}_2$  and be consistency of using “ $\text{VPDB-CO}_2$ ” throughout the paper. The same changes should be made for the axes in Fig.3 and Fig. 4.*

A: Both notations are in frequent use. There is no ambiguity in either form nor is there any rule as to which form should be preferred.