The authors would like to thank both Reviewers and the Editor for their time and effort in reviewing our manuscript entitled "QBOi El Niño Southern Oscillation experiments: Teleconnections of the QBO". Above all, the authors are deeply grateful for the many insights gained by reading the papers recommended by the reviewers. We will incorporate their valuable comments and suggestions on how our proposed revised manuscript will address their concerns. Our reviewer responses and revision plan are shown in blue text whereas reviewers' and Editor's comments are shown in black. Individual responses to the Editor are as follows.

Black: Editor's comments

Blue: Authors response to the Editor

To Editor:

EC1: 'Editor Comment on egusphere-2025-1148', David Battisti, 17 May 2025

Dear Dr. Naoe and colleagues:

Both Anonymous referees have posted their comments on your manuscript (WCD 2024-1148). As per WCD policy, you are now to post a response on how you will address the referee's comments — after which I will make a decision on the manuscript. Both reviewers have made excellent comments on the manuscript and call for revisions (one major, one minor). To provide guidance in revising the manuscript so that it is acceptable for publication in WCD, below I itemize the issues that I expect will be addressed in a revised manuscript. I will also post these on the WCD page for the manuscript.

Both anonymous reviewers feel this is a worthwhile manuscript for publication in WCD, and I agree.

The opening paragraph by Reviewer #1 has a very succinct summary of the paper, followed by 8 bulleted points that contain either comments or suggestions. I strongly recommend you address all the comments, and adopt all the suggestions. In particular, the reviewer notes that the text is not sufficiently critical of the model results concerning the impact of the QBO on the polar vortex (Figs. 1 and 2), stating: "only ECCAM5, WACCM and MRI are reasonably correct for neutral ENSO, but none get El Nino right. Maybe ECCAM5 and MRI get LaNina right (relative to ERA5)." I agree: only two of the models get within 1/2 the amplitude of the observed for the ENSO neutral case (MRI and WACCM), and only the MRI model also shows a stronger impact on the QBO on the vortex under La Nina conditions than under El Nino conditions (but even that model has the wrong sign response for El Nino conditions). The reviewer also asks for more clarification on the text on lines 221-226, and clarification on the statistical significance when multiple indices are used in the identification of the QBO. Reviewer #1's suggestion "to include more in-text references to figure panels being discussed" would really help the reader.

- In particular, the reviewer notes that the text is not sufficiently critical of the model results concerning the impact of the QBO on the polar vortex (Figs. 1 and 2), stating: "only ECCAM5, WACCM and MRI are reasonably correct for neutral ENSO, but none get El Nino right. Maybe ECCAM5 and MRI get LaNina right (relative to ERA5)." I agree: only two of the models get within ? the amplitude of the observed for the ENSO neutral case (MRI and WACCM), and only the MRI model also shows a stronger impact on the QBO on the vortex under La Nina conditions than under El Nino conditions (but even that model has the wrong sign response for El Nino conditions).

Thank you very much for your suggestions to note that the text is not sufficiently critical of the model results concerning the impact of the QBO on the polar vortex (Figs. 1 and 2). We will revise the text to add these points and delete unnecessary descriptions.

- The reviewer also asks for more clarification on the text on lines 221-226, and clarification on the statistical significance when multiple indices are used in the identification of the QBO.

The analysis in the Walker Section initially used a consistent QBO definition and target season across all models. Specifically, we define the QBO using the zonal-mean zonal wind at 70 hPa during JJA. The corresponding figures are now included in Fig. R2-3 of the responses to Reviewer #2 and supplementary material. With this uniform framework, we identify a coherent signal, but we want to enhance this signal and capture the strongest response in each model. To do so, we allow for slight adjustments in the QBO definition (70 or 85 hPa and JJA or SON) and in the target season (ranging from May to November), when necessary. We will clarify this process in the revised text of Section 5.2. Importantly, when a model's QBO definition differs from the standard (70 hPa during JJA), we account for the increased flexibility by applying a Bonferroni correction. This reduces the significance threshold (alpha) from 0.05 to 0.025 or lower, depending on the number of alternative definitions tested. The significance threshold (also called the significance level) determines the p-value below which the null hypothesis is rejected. If the p-value is smaller than α, we reject the null hypothesis and consider the result statistically significant. We will clarify this point in the revised version of Section 2.

- Reviewer #1's suggestion "to include more in-text references to figure panels being discussed" would really help the reader.

Thank you very much for the suggestion from Reviewer #1 and the Editor. We will revise the text to include more in-text references to figure panels being discussed.

Reviewer #2 also has excellent major comments and I strongly recommend you address them in your revised manuscript. In particular, Reviewer #2 asks for more discussion and analysis of why almost all the models do not reproduce three of the four teleconnections examined, and I agree. In some cases, further analysis may be required to support these discussions (e.g., is there a relationship between the model biases in the strength of the simulated QBO (in either neutral, El Nino and La Nina conditions) and the strength of the polar vortex response? Is there a relationship between the model biases in the strength of the polar vortex and the polar vortex response? Is there a relationship between biases in the extratropical stratospheric winds and the weakness in the impact of the QBO phase on the polar vortex?).

We appreciate your helpful suggestions. In the revised manuscript, we will include more discussion and analysis of why almost all the models do not reproduce teleconnections examined. Please see our response to Reviewer #2 major comment R2-1 in more detail. Here, a summary of this discussion is as follows:

QBOi ENSO experiments

ENSO modulation of the QBO in our QBOi ENSO experiments is investigated by a core paper of Kawatani et al. (in revision).

https://egusphere.copernicus.org/preprints/2024/egusphere-2024-3270/

- QBOs in some models are irregular, from a simple, time-height cross-section of the monthly and zonal winds in the El Nino and La Nina simulations, as shown in Figure 2 of Kawatani et al.

a) QBO teleconnections to polar vortex

- Problems of QBO teleconnection to the stratospheric polar vortex were investigated in detail by previous studies (Bushell et al., 2022; Anstey et al., 2022). As Anstey et al. (2022) described, the strength of the QBO teleconnection to the NH winter stratospheric polar vortex was shown to correlate with the amplitude of the QBO at 50 hPa. This altitude is the strongest correlation with the vortex in observations.

- Most models show poor performance of QBO amplitude at 50 hPa while climatological polar vortices in NH winter can be reproduced with their strength. These results are consistent with the hypothesis that unrealistically weak low-level QBO amplitude can weaken the teleconnection.

b) QBO teleconnections to subtropical jet

- Models with larger QBO amplitudes do not necessarily exhibit stronger jet responses, nor do models with smaller amplitudes consistently show weaker responses. This means that neither the QBO amplitude nor the APJ position explains the inter-model spread in the QBO-APJ connection. Other factors may determine the QBO-APJ connection in the model.

c) QBO teleconnections to tropical precipitation

- The combination of stratospheric and tropospheric biases in the tropics weakens the QBO signal reaching the tropical troposphere and contributes to inter-model differences in both the timing and spatial manifestation of the teleconnection.

Reviewer #2 also notes that previous work suggested that a measure of the efficacy of a model to reproduce QBO's impact on the polar vortex (the Holton-Tan effect) seen in observations is sensitive to the level that is used as an index of the QBO, and that model differences in the QBO justify the use of model-specific indices. Please address this point in your revised manuscript. Also, if you did choose levels to define the QBO that were model-specific, would the QBOs simulated by the models still be only half as strong as that observed (as documented in Fig. 3)? Would that still be the leading candidate for the weak relationships between the QBO phase, ENSO phase and polar vortex?

We appreciate your helpful suggestions. As we already described before, the problems of QBO teleconnection to the stratospheric polar vortex were investigated in the previous studies in detail. The QBO teleconnection to the NH winter stratospheric polar vortex is the strongest in observations when the QBO index is taken at 50 hPa (Anstey et al. 2022). Thus, we want to do model-observation comparison by applying the same QBO phase definitions to the models that are optimal for observed teleconnections, in order to determine if observed teleconnections are present in the model runs, without adjusting them on a model-by-model basis, for all analyses presented in this article.

In order to answer Reviewer #2 and Editor questions, we check levels to define the QBO that are based on observational studies (i.e., at 50 hPa) and that are based from a model specific level (i.e., at 30 hPa), as shown in Fig. R2-4 of the responses to Reviewer #2. Both panels (QBO50 and QBO30) show that most models underestimate QBOs and they are struggling to reproduce observed polar vortex responses to the QBO. Please see our response to Reviewer #2 major comment R2-4 in more detail.

Finally, Reviewer #2 asked why a different analysis procedure was used to examine the relationship between the QBO phase and the Walker circulation than that used to examine the other three teleconnections and whether the teleconnections were stronger for the Walker circulation simply because optimal pressure levels and seasons were chosen. I am not to bothered by this because, to be frank, the evidence presented in this section is pretty damning. Contrary to the description in the text, the observed relationship between the Walker circulation and the phase of the QBO shown is not well reproduced by most of the models for either La Nina or El Nino conditions. For La Nina conditions (Fig. 11), the anomalies in the zonal winds over the Pacific show a slightly westward shifted Walker circulation, whereas the models b,d,g,h and i shows a weakened Walker circulation (in phase anomalies of the opposite sign as the climatology aloft) and model e shows only easterly anomalies everywhere. The agreement during El Nino conditions is even worse (Fig. 12). [By the way, please note the contour interval for the anomalies in these figures. They seem to be much coarser than the discretized colorbars.] Stepping back a bit, I wonder whether the relationship between the QBO phase and the Walker circulation is poor because the band to define (5S-5N) the

Walker circulation may be too narrow; 10S to 10N would better capture the zonal wind anomalies associated with the Walker circulation. Based on Fig. 17.17 of Wallace et al (2023), I expect this isn't the answer -- but it might be worth checking.

Thanks for your comments and suggestions. In response, we will revise the analysis using the 10ºS-10ºN band, which better captures the zonal wind anomalies. The updated main figures now reflect this broader latitude band. Additionally, to provide more context and clarity, we will include the results from our initial analysis, which focused on the target season JJA and used the standard QBO definition (zonal-mean zonal wind at 70 hPa during JJA) in the supplementary material. One of these figures (LN experiment) is presented in Figs. R1-1 (and R2-3; both figures are the same) of the responses to the reviewers. Such supplementary figures allow readers to better understand the progression of our approach. We also slightly adjust the main figures to align more closely with the standard QBO definition and the JJA season. The manuscript text will be revised accordingly to enhance clarity and ensure that the description of model performance is accurate.

Minor comments:

Does the GPCP bar in panel 9b stop the top of the plot, or does it run off scale? Why isn't there an error bar on this bar?

The GPCP bar runs off scale, so much so that even the error bar doesn't appear. The reason, to some extent, for this is the large signal due to the QBO-ENSO aliasing the manuscript discusses. We have produced two sets of figures for this plot, one where the y axis limits are set based on the GPCP bar and the other, like the original, where the limits are fixed to make the plot clearer. Both have positive and negative factors, and we will provide the full figure in the revised Supplementary section.

In Fig. 10, is the temperature also the zonal mean over the western Equatorial Pacific, or is it a zonal mean?

The temperature is also the mean over the western equatorial Pacific only. Thank you for the question; the revised manuscript clarifies this issue.

Lines 777-790: These statements are inconsistent with the published papers, dating back as far as Hoerling et al. (1987). Atmosphere general circulation models DO robustly reproduce the nonlinearity in the atmospheric response to warm and cold ENSO phases, given El Nino and La Nina SST anomalies.

Thank you for your suggestions. Our understanding is that there exists observational evidence of mutual interactions between ENSO and QBO, but this possibility has not been widely studied using CMIP-class, climate model simulations. The observed ENSO/QBO relationship in current climate models is generally poorly reproduced, likely as a consequence of the coarse spatial resolution and the reliance on stationary parameterizations.

Serva, F., Cagnazzo, C., Christiansen, B., Yang, S.: The influence of ENSO events on the stratospheric QBO in a multi-model ensemble, Clim. Dyn., 54, 2561-2575, 2020, https://doi.org/10.1007/s00382-020-05131-7

Also, given the very weak relationship between the QBO phase, the ENSO phase, and the tropical anomalies shown in this study, it is unlikely that weaker ENSO events or ENSO events with less dramatic changes in the location of tropical convection than used in this study would yield further insights.

We agree with your second point that it is unlikely that weaker ENSO events or ENSO events with less dramatic changes in the location of tropical convection because of a weak relationship between the QBO phase, the ENSO phase, and the tropical anomalies are shown in this study. But, one study indicated that QBO is also influenced by the tropical SSTs in the Central Pacific (Shibata and Naoe, 2022), so that we believe that it will be worth further study of the role of ENSO flavors in the QBO-ENSO teleconnection.

Shibata K, Naoe H, 2022: Decadal amplitude modulations of the stratospheric quasi-biennial oscillation, J. Meteorol. Soc. Japan, 100, 29-44, https://doi.org/10.2151/jmsj.2022-001