The revised manuscript was kindly assessed by one of the original reviewers, who is now largely satisfied but notes several areas requiring final attention -- including clarifying key methodological details and justifying specific choices concerning the hypotheses and analysis plans. We are now much closer to in-principle acceptance (IPA), so provided you are able to respond comprehensively to these points in a final revision and response, IPA should be forthcoming without requiring further in-depth Stage 1 review.

**We extend our many thanks to the reviewer for the invaluable insight into improving the submission from the very first up to this third round of revision. We address each point below individually. These revisions have undoubtedly improved the clarity of our proposed research. Our responses to the reviewer's comments are marked in 'bold" and the new added text to the manuscript is marked in 'green'.**

Many thanks to the authors for their responses and clarifications. The rationale for conducting this study using the 22-item RWA scale is now clear to me, and I agree that it is worthwhile to test the predictions that arise from system justification theory and the dual process model, which conceptualise right-wing authoritarianism as a unitary construct. I think the authors make an important point when they state, in their response to my comment, that "**This is, of course, a direction for future research that is worth exploring, particularly whether or not the three separate subscales map better at a neural level than the superordinate RWA scale.**" I suggest that this point should be included in the manuscript.

**We thank the reviewer for this recommendation and fully agree that this should be explicated in the manuscript and thus, included this in the manuscript (p. 16):**

"**RWA as a multidimensional construct is a direction for future research that is worth exploring, particularly whether or not the three separate subscales map better at a neural level than the superordinate RWA scale.**"

The explanation of the predictions of overlapping and distinct brain regions is clear now.

There are two points in the section added to page 16 of the manuscript that I think require some clarification – see table below. I also have some feedback on the Whole Brain Analysis, and two additional queries about the specification of the ROIs (see table below). I apologise that I did not raise these queries about the ROI specifications in Round 2.

| Page | Line(s) | Comments |
|------|---------|----------|
| 16 | 11 | exploring RWA as a stable unidimensional trait<br><br>This reads as ambiguous to me. A more precise phrasing might be something along the lines of "exploring whether RWA – conceptualised as a unidimensional trait – is reflected in brain structure". |

| | | We thank the reviewer for highlighting this. We agree that this way of phrasing is much more clear and have edited the manuscript accordingly (p. 16):<br><br>"**Nevertheless, we have good reason to believe that there is merit in exploring whether RWA – conceptualised as a unidimensional trait – is reflected in brain structure.**" |
|---|---|---|
| 16 | 19-22 | Moreover, one study conducted a set of factor analyses of RWA (and SDO) showing that both multidimensional and unidimensional models of RWA demonstrate acceptable fit to response data granted the items themselves were already divided into their respective subscales (Kandler et al., 2016).<br><br>This is quite difficult to follow, especially the statement, "granted the items themselves were already divided into their respective subscales". Was this a factor analysis of the same 22-item RWA scale used in the present study? In which case, as the authors have pointed out, factor analysis would not be appropriate or informative, given the double- and triple-barrelled nature of several items. It would also be helpful to specify (a) if the multidimensional model(s) that demonstrated acceptable fit was a 3-dimension model, and if more than one multidimensional model demonstrated acceptable fit; and (b) whether the multidimensional model(s) or the unidimensional model demonstrated better fit to the data. I think it would also be appropriate to specify the number of participants whose data were used in Kandler et al.'s factor analysis.<br><br>**We thank the reviewer for seeking clarification on this. The Kandler et al. study did not use the 22-item RWA scale. They opted for the 12-item variation of the RWA scale which can be clearly delineated into the three subscales (i.e. no double/triple-barelled items). We believe this further supports our stance in that even after the removal of items that overlap across the three covarying traits, the factor analyses conducted demonstrated that RWA (and SDO) may be best conceptualised using "the more parsimonious unidimensional model". They also added that "modelling RWA and SDO as multidimensional constructs did not provide an appreciably better model fit".**<br><br>**We agree that specifying which multidimensional models showed acceptable fit as well as the number of participants used in that study would be appropriate to include in the manuscript (p. 16):**<br><br>"**Kandler et al. used a 12-item variation of the RWA scale with three clearly delineated subscales (and the original 16-item SDO scale with two subscales). Recruiting a large sample (N = 1437), the responses of these individuals to the RWA and SDO scales were subjected to confirmatory factor analyses fitting three alternative models. Their findings indicated acceptable fit to the data where RWA was modelled as three factors, one factor and one superordinate factor with three subordinate factors. Overall, the authors concluded that the multidimensional models do not provide substantially better fit that warrants opting out of a more parsimonious unidimensional conceptualisation of RWA (and SDO). We believe this further supports our stance in that even after the removal of items that overlap across the three covarying traits, RWA may be best conceptualised as a unitary construct.**" |
| 19 | 12 | We intend to measure the mean grey matter volume (GMV).<br><br>I'm not sure what is meant by 'mean' here. As I understand it, the authors will conduct a standard mass univariate multiple regression VBM analysis. Usually, |

| | | |
|---|---|---|
| | | whole-brain interrogation of the results would proceed thus: Once the GLM has been fitted, one uses the information from the resulting beta and ResMS images (as well as the SPM.mat file) to generate images of t statistics. Contrast vectors are specified, which indicate the linear combination of beta images to test. These linear combinations are then used to identify any regions (clusters of voxels) in which the contrast is significantly different from zero.<br><br>**We thank the reviewer for pointing this out. We fully agree with the reviewer in that 'mean GMV' is not an appropriate characterisation of the VBM analysis the reviewer correctly described above. Therefore, we have removed this line.** |
| 19 | 14 | It is true that the "factorial design" (including multiple regression) models in SPM use ordinary least squares, but this is a confusing way of explaining the model that will be applied. Note that SPM applies the regression model in every voxel (mass univariate analysis), so it is not mean grey matter volume that is the dependent variable in each multiple regression analysis, but smoothed, corrected grey matter volume estimates for each participants for that voxel.<br><br>**We thank the reviewer for this suggestion and have edited the manuscript appropriately (p. 19):**<br><br>**"Each multiple regression analysis will use ordinary least squares models with smoothed, corrected GMV estimate as the dependent variable, and RWA (or SDO) score, gender, age and TIV as independent variables."** |
| 21 | 7 | What is the justification for averaging the mean volumes from the left and right amygdala? Why not conduct a region of interest VBM using a bilateral amygdala mask? As it might be the case that only left or right amygdala is associated with RWA/SDO, or that RWA/SDO scores are only associated with regional grey matter volume in a part of the 20mm spheres used as amygdala ROIs.<br><br>**We thank the reviewer for this clarification. We do predict that RWA and SDO will correlate with both left and right amygdala volumes. Importantly, we selected the spherical ROI specifically to determine whether the same region in the amygdala that is associated with system justification scale score (Nam et al., 2017) would also be associated with system-justifying ideologies, namely RWA and SDO. The averaging of left and right amygdala volumes was meant to be used to show how the bilateral amygdala volume correlate with RWA/SDO in a scatter plot graph of regional volume against RWA/SDO as with Nam et al (2017). We intend to do the same for all other ROIs. However, we can understand how this can be misconstrued as the mean volume being used in the ROI analysis. We have edited the manuscript to clarify this ambiguity (p. 19):**<br><br>**"After the ROI analysis, we will conduct Pearson's correlation tests for each ROI GMV with RWA and SDO separately. With regards to amygdala GMV, we average the volumes from the left and right amygdala to generate a single measure of bilateral amygdala volume."** |
| 21 | 10 | Why is the region of interest analysis confined to the left insula? Hypothesis 4 does not specify left insula, which I took to mean that the hypothesis was bilateral. |

| | | We thank the reviewer for highlighting this important point. Cazzato et al and Chiao et al both specified SDO to be associated with the left anterior insula (among other regions). Therefore, all mention in the manuscript relating to H4 were edited to specify the left anterior insula (consequently, these will be highlighted in 'green'. |