08/12/2021

Dear Prof Chambers,

Neuroanatomical Correlates of System-justifying Ideologies: A Pre-registered Voxel-based Morphometry Study on Right-Wing Authoritarianism and Social Dominance Orientation

Jan Paolo M. Balagtas, Serenella Tolomeo, Bindiya L. Ragunath, Paola Rigo, Marc H. Bornstein & Gianluca Esposito

Three expert reviewers have now evaluated the Stage 1 manuscript. All find the work to address a valid scientific question; however, as you will see, the reviews raise a number of major concerns that preclude Stage 1 in-principle acceptance (IPA) in its current form. Key shortcomings include lack of clarity and justification of the hypotheses, and the lack of suitability of the design for answering the research question -- including, especially, the use of an ROI-based approach instead of a whole-brain or combined approach and the validity of the behavioural measures for assessing system justification. The reviewers also query the logical coherence of the underlying rationale, the extent to which the design accommodates sampling characteristics, and a range of areas lacking in methodological detail such as the procedure for defining the ROIs.

For a regular manuscript, a set of reviews this critical would lead to outright rejection, but the advantage of the Registered Reports process is that it offers authors the possibility of resolving design limitations before they become roadblocks. In this case, given these assessments, I am happy to invite a Major Revision but do so with the caveat that a substantial amount of work will be needed to achieve IPA, and all criticisms must be comprehensively addressed.

We very much appreciate the opportunity to resubmit our work. We are grateful for the points raised by reviewers which we address in a point -by-point fashion below. In summary, we followed the reviewers' suggestions to improve clarity of the rational and methodological design as well as justification of the hypotheses. We believe all these changes strengthen our paper substantially. Thanks again to the reviewer for pointing us to these issues and for suggesting how to address them. Our responses to the reviewer's comments are marked in 'bold' and the new added text to the manuscript is marked in 'green' also in the respective document.

Reviews

Thank you for the opportunity to review this proposed analysis. I think it is an interesting and worthwhile extension of prior neuroanatomical investigations of social and political attitudes that has the potential to expand our understanding of the universality and/or boundaries of the neurobiological bases of ideologies by examining these relationships in the Singaporean context. The authors articulate clear hypotheses and a largely clear analysis plan. I hope they will consider my comments as they continue with their work.

Q1) Although I understand the reasoning of treating system justification theory as an umbrella for authoritarianism and social dominance orientation, I was a bit surprised (given the introductory theorizing) that the authors do not have a plan to measure system justification more directly. Insofar as the authors have an interest in SJT and consider RWA and SDO to be system-justifying ideologies, I think it would make sense to measure the 8-item SJ scale (Kay and Jost 2003) so that the authors can (a) examine the predictive value of RWA and SDO above and beyond SJ per se, and (b) assess whether the Nam et al. (2018) findings of a correlation between amygdala structure and SJ can be replicated in the Singaporean sociopolitical context. I suppose this is in some sense a partial registered report in that the data have already been collected, so perhaps there is not much room for my suggestion here. But if the authors have the ability to re-contact participants for another short survey, I would suggest that they administer SJ to consider in their analyses. Otherwise, I'm afraid that the SJT framing of the introduction is a bit misleading.

R1) We thank the reviewer for this suggestion. Since our data have already been collected, we have no opportunity to re-contact participants for another short survey. However, we agree with the reviewer and we have now removed the misleading text regarding system justification theory in the introduction and added the following statement section (p.4):

‘’Jost & Hunyady (2005) argue that individuals adopt different system-justifying ideologies as a means to make sense of the status quo. These include, but are not limited to, meritocratic ideology, political conservatism, belief in a just world, RWA and SDO. These ideologies serve the same purpose of rationalising and legitimising social inequalities but the observable manifestations of each ideology may differ.”

Q2) It may also be worth noting that Nam et al. (2018) did explore correlations between SDO and brain structure in a subset of their sample, but did not find significant associations (see Supplementary Discussion). The authors may want to take a look at this as further context for their hypotheses, especially considering my above point and that they have a more well-powered sample to assess the relationship.

R2) We agree that Nam et al. (2018) explored the relationship between SDO and brain structure, however, our hypotheses aimed to address RWA and SDO and their brain structure correlates with a more well-powered sample. To make this more explicit, we have added more discussions in the introduction (p.17):

‘’It is worth noting that Nam et al. (2017) previously did not find a significant correlation between SDO and amygdala volume. However, the absence of a relationship may likely have been due to the relatively small sample size (N = 37) used to analyse this relationship. Equipped with a more well-powered sample, our study (N = 82) presents a more definitive measure of a neuroanatomical correlate not only of SDO but also RWA.”

We hope the reviewer agrees that we reported our power analysis on p.16 and our study is well powered to address our hypotheses. We have clarified in the methods on p.16:

“Our study exceeds both criteria with a final sample size of N = 82, and consequently, we believe this to be a well-powered sample to provide an accurate estimate of the effect size regarding system-justifying ideologies and their respective structural brain associations.”

Q3) It would be helpful to get some more explanation about the sociopolitical context of Singapore. For example, does political ideology map onto a left-right dimension as it is often conceptualized and measured in Western contexts? I think it would make sense to explain the nature of more explicitly political ideology in Singapore, and how RWA and SDO might relate to this construct. For example, are the authors theorizing, in some sense, that RWA and SDO comprise dimensions of political orientation that can be mapped onto both shared and distinct brain regions?

R3) We concur fully with the reviewer, however we did not have a specific hypothesis on how left-right dimension relates to both shared and distinct brain regions so we have mainly provided more explanation about the sociopolitical context of Singapore.

In addition, we are grateful to the reviewer for providing us with such a crucial suggestion on how to explicitly explain the political ideology in Singapore and how it relates to RWA and SDO. We have added this new text on p.7:

“To reiterate, the correlation of RWA and SDO is largely a function of the ideological contrast of the national context it is embedded in (Rocatto & Ricolfi, 2005). Specifically, countries that have a strong ideological contrast tend to promote a political left-right dimension such that individuals that ascribe to a specific political ideology (i.e. left vs right) tend to show positively correlated RWA and SDO scores. More relevant to this paper, the sociopolitical context of Singapore does

not feature an explicit left-right dimensionality and so, would be considered a country of low ideological contrast. Therefore, the DPM model would predict that RWA and SDO would be largely independent of each other in the Singapore context."

We are also grateful to the reviewer for highlighting an important query with regards RWA and SDO comprising dimensions of political orientation that can be mapped onto brain regions. As such, we would like to address it directly here in this response letter:

Although the correlation between RWA and SDO scores vary as a function of the degree of ideological contrast of the population of study, this does not preclude the distinction of antecedents of RWA and SDO as predicted by the DPM model. That is, RWA is largely motivated by a dangerous worldview whereas SDO is largely motivated by a competitive worldview. In addition, the amygdala is also involved as this region is implicated with legitimising social hierarchies, both promoted by RWA and SDO ideologies. In other words, the degree of ideological contrast has more to do with the correlation between RWA and SDO and much less the underlying psychological and neuroanatomical mechanisms underlying them. Theoretically, regardless of the degree of ideological contrast of a given population, the shared and distinct brain regions of RWA and SDO would remain stable. That being said, a cross-national comparison of RWA and SDO scores and their neuroanatomical correlates between countries of different ideological contrasts must be conducted to directly address this question. In which case, our current study design is not equipped to tackle such a research question and hence, we are not inclined to make any specific hypotheses in our manuscript.

Q4) My understanding is that ethnic Chinese are the dominant ethnic group in Singapore. How does a sample composed primarily of dominant group members affect the authors' interpretations of the correlation between brain structure and RWA/SDO? I'm wondering this especially given the theorizing regarding the Singaporean context of institutionalized multiculturalism.

R4) We very much agree with this comment and we have now emphasised this point in the new text (p.12):

''It is worth noting that the sample represented in this study is only of the majority Chinese ethnic group. In a review of the influence of policies on intergroup relations, Guimond et al. (2014) assert that the impacts of intergroup-related policies, such as multiculturalism policies, vary to the extent that cultural perceptions of majority group individuals towards such policies may not represent those of minority group individuals. Therefore, findings in this paper may not extrapolate to the entire Singapore population and may be restricted to those of the ethnic majority group. Nonetheless, the study still holds merit in that limiting the study to only the majority ethnic group can allow the focus on a relatively homogenous sample to maximise statistical power. Future studies that examine neuroanatomical correlates of RWA and SDO scores should subsequently include ethnic minority groups to examine possible differences in their neuroanatomical correlates of RWA and SDO."

Q5) How was the 22-item subset of the RWA scale chosen from the original 32 items? Some justification for the subset would be helpful.

R5) We have added the required justification on p.14:

''The RWA scale in this study used a 22-item version (Altemeyer et al., 2006; Appendix A). The 22-item scale is used in this study as it is the most updated version of the RWA scale and demonstrates comparable psychometric properties as the original 32-item scale. Therefore, we opted for this shortened version of the scale."

Q6) Similar prior work (Kanai, Feilden, Firth, and Rees 2011; Nam et al. 2018) used a Gaussian kernel of 12 mm FWHM as part of the image preprocessing. Could the authors give some explanation of their reasoning for using a different metric (8 mm)?

R6) We agree with the reviewer and we have now decided we want to be consistent with prior work by Kanai et al., 2011 and Nam et al., 2018 so we have now edited the Gaussian Kernel of 12mm FWHM on p.17:

‘’Finally, they will be smoothed by convolving the images with an isotropic Gaussian kernel of 12 mm full width at half maximum (FWHM; Ashburner & Friston, 2005; Kanai et al., 2011; Nam et al., 2018).”

Q7) How will ROIs be determined? Using ROI masks? Specifying volumes around a set of coordinates?

R7) We address these point in the methods on p.18:

‘’We will investigate the association between GMV and scores in the RWA and SDO scales using ROI multiple regression analyses. As the brain areas in our hypotheses are identified with a strong a priori prediction, the threshold of significance was set at $p < 0.05$, with small volume correction for multiple comparisons in the ROIs. To analyse these regression models, we will be using the MarsBaR toolbox (http://marsbar.sourceforge.net/) to anatomically define ROIs for the three pre-specified brain regions (i.e. amygdala, vmPFC and insula) according to our main hypotheses. The GMVs will be extracted from their ROIs using anatomically defined spheres with a radius of 20mm centered at (MNI: x=−36, y=−9, z=−17) for the left amygdala and at (MNI: x=27, y=12, z=−21) for the right amygdala (Nam et al., 2017). We then average the mean volumes from the left and right amygdala. The ROI of the vmPFC will be anatomically defined as a sphere with a radius of 20 mm centered at (MNI: x = 0, y = 40, z = -18; Li et al., 2017). Finally, the ROI of the insula will be anatomically defined as a sphere with a radius of 20 mm centered at (MNI: x = -45, y = 26, z = -6; Chiao et al., 2009; Cazzato et al., 2015).”

Q8) It’s a good idea to include age as a covariate in the model. I would also suggest gender/sex and global brain volume as baseline covariates (following previous work).

R8) We agree with the reviewer and we will include age, gender and global brain volume as a covariate in the model. Please see the detailed description on p.19:

‘’In these analyses, we intend to control for total intracranial volume (TIV), age and gender by including them into the regression model as independent “nuisance” variables. TIV is an important variable to account for particularly in ROI-based volumetric measures because such subtle differences in regional brain volume may be confounded by individual differences in overall brain size (O’Brien et al., 2011). We are also controlling for age not only because TIV varies as a function of age (Bartholomeusz et al., 2002), but also because both RWA and SDO have been shown to decrease with age (Altemeyer, 1998; Ruffman et al., 2020). Accounting for age is also necessitated in this study because the analysis will include participants from two different age groups, a young adult sample and a middle-aged adult sample. We would expect both self-report and volumetric brain differences between these two age groups so including age in the regression model will minimise confounds due to age differences. Finally, past research also suggests a gender difference in self-reports of RWA and SDO. In particular, women tend to report higher RWA scores than men (Brandt & Henry, 2012) whereas men tend to report higher SDO scores than women (Pratto et al., 1994). Combined with an overall brain volume differences between men and women (Kaufmann et al., 2001; Ruigrok et al., 2014; Takahashi et al., 2011), we reckoned to control for gender would facilitate in identifying significant neuroanatomical correlates, as we predict with the age variable. We would like to emphasise that although system-justifying ideologies and regional (and overall) brain volume do seem to vary with age and gender, these are treated as nuisance variables in the main analysis as they do not comprise the main objectives of the study.”

Reviewer 2

The authors propose to investigate correlations between regional grey matter volumes and right-wing authoritarianism (RWA) and social dominance orientation (SDO). The research question is clearly defined and makes sense in light of theory. In general, the authors did a fine job in presenting their work. I have a few comments on specific parts of the manuscript, which are detailed in the table below. I have one concern about the proposed work plan, but I think it is easily remedied. My concern is that the proposed plan is to conduct ROI analyses to answer questions that can only be answered by whole-brain analysis. The questions are, 1. Is/are there any overlapping brain region/s that are related to RWA and SDO? 2. Is/are there any region/s that are related to RWA but not SDO? 3. Is/are there any region/s that are related to SDO but not RWA? A whole-brain analysis is needed to answer these questions. However, given that the region-specific hypotheses are pre-registered, it would be fine to also conduct these ROI analyses with small-volume corrections, provided that the way the ROIs will be defined is fully specified in the Stage 1 manuscript. In case it is helpful, here is a link to a previous pre-registered VBM study in which a whole-brain analysis was conducted in addition to ROI analyses: The relationship between individual differences in gray matter volume and religiosity and mystical experiences: A preregistered voxel-based morphometry study - Elk - 2020 - European Journal of Neuroscience - Wiley Online Library

Page	Line(s)	Comments
3	**5**	I'm not clear what 'hallmark characteristic' means in this context. Is authoritarian submission a core component of RWA ideology, such that if one doesn't exhibit authoritarian submission, then one can't be categorised as holding RWA ideology? Or is it possible (albeit unusual) to hold RWA ideology but not be obedient to authority? I'd recommend re-phrasing this sentence to clarify. **We thank the reviewer and have made the necessary edit to clarify authoritarian submission as a covarying trait (with the other two traits) of RWA ideology (p.3):** **"Altemeyer (1998) conceptualised right wing authoritarianism (RWA) as an ideology that can be understood as a cluster of three covarying traits: authoritarian submission, authoritarian aggression and conventionalism. That is, these traits comprise a singular measure of RWA. Authoritarian submission or the tendency to almost unquestioningly obey an authority figure is one such hallmark trait of RWA."**
3	**6**	The word 'unquestionably' -- should this be 'unquestioningly' ? **We thank the reviewer and have incorporated the recommendation in the above point.**

3	**26**	"*authoritarian submission, and by extension the RWA ideology*" – I think there is a logical fallacy here. Even if those who hold RWA ideology always exhibit authoritarian submission, it does not mean that those who exhibit authoritarian submission will always hold RWA. **We thank the reviewer for this point and we would like to further clarify on how the three traits are related to RWA. The original conceptualisation of the RWA scale is such that theoretically, the three traits are subsumed under a unidimensional construct of RWA, not only because they are correlated but they covary with one another. This is likely to be a result of how the items were phrased. Visual inspection of the scale would reveal that a handful of the items are double or triple-barrelled, in that they measure more than one of the covarying traits. To illustrate, one of the items can be divided into their individual traits: "Our country desperately needs a mighty leader (authoritarian submission); who will do what has to be done to destroy (authoritarian aggression); the radical new ways and sinfulness that are ruining us (conventionalism)". Nevertheless, both the 32-item and 22-item versions of the scale demonstrate acceptable psychometric properties and are the most widely used version of the RWA scales and were therefore utilised in this manuscript. In this way, the logic of authoritarian submission having a biological basis would extend to the RWA ideology, which has also been noted in other published work (e.g. Warner, Tranel & Asp (2016) The Henchman's Brain Neuropsychological Implications of Authoritarianism and Prejudice).**
8	**16-18**	This argument seems to undermine the rationale for the study. If self-report measures are insufficient or inaccurate measures of RWA and SDO, then how does it help to look at the neuroanatomical correlates of scores on these self-report measures? **We thank the reviewer for this suggestion and have made the necessary amendment in the introduction (p.8):** **"Although traditional self-report measurements of RWA and SDO have demonstrated robust reliability and validity across multiple studies, the examination into the neural bases of RWA and SDO can provide more solid evidence for their status as stable individual differences."**
8	**27-32**	This is a little unclear. To me, it reads as though the two predictions are opposing, whereas in fact, they are concurrent predictions. **We thank the reviewer for this highlighting this point and have made the necessary amendment in the introduction (p.8):** **"We predict that RWA and SDO would involve identical brain regions as they are both system-justifying ideologies that individuals espouse to maintain the hierarchical structure of society. Additionally, these constructs correlate but are nonetheless independent, and would therefore recruit unique brain regions to differentially substantiate**

		these ideologies in terms of antecedents and outcomes as propounded by the DPM model."
9	**12**	The Nam reference here lists 2017 as the publication date, but the entry in the references list says 2018. **We thank the reviewer for noticing and have made the necessary amendment in the reference list (p.28):** **"Nam, H.H., Jost, J.T., Kaggen, L., Campbell-Meiklejohn, D., Van Bavel, J.J., 2017. Amygdala structure and the tendency to regard the social system as legitimate and desirable. Nat. Hum. Behav. 2 (2), 133-138. 10.1038/s41562-017-0248-5."**
9	**14**	The directionality of the hypothesised correlation should be included. **We thank the reviewer for this suggestion and have made the necessary amendment in the introduction (p.9):** **"and hence, the scores in these measures will presumably also correlate positively with amygdala volume (H2)."**
9	**31**	It's unclear if the study being described here is another vmPFC ablation study, or a different type of study. **We thank the reviewer for bringing up this clarification and have made the necessary amendment in the introduction (p.10):** **"Importantly, after partialling out religious fundamentalism, Asp, Ramachandran & Tranel (2012) found that the degree of vmPFC damage still led to higher RWA scores, suggesting that the fundamentalist beliefs associated with RWA is not exclusive to a religious context (Krauss, Streib, Keller & Silver, 2006)."**
10	**1**	The use of the word 'poor' here reflects a value judgement and could be misunderstood by readers who consider authoritarian submission and gender stereotyping to be good things. **We agree with the reviewer and have made the necessary amendment in the introduction (p.10):** **"In other studies, vmPFC lesions are also implicated in changes in performances of tasks related to the covariations of RWA"**
10	**5-9**	It's not clear from this what Grafman's findings were, so it's difficult to follow the logic of the subsequent inferences. **We thank the reviewer for this point and have decided to remove this text to minimise confusion.**

10	**18**	It's not clear what seemingly inconsistent findings are being referred to here. **We thank the reviewer for this point and we have deemphasised the inconsistencies and focused on variation in brain areas associated with SDO across fMRI studies this in the main text (p.10):** **"These variations in brain areas are potentially attributable to the different tasks employed in each study."**
10	**20**	Was it higher or lower SDO scores that were associated with reduced ACC & insula activation? **We thank the reviewer for highlighting this point and we have made the necessary amendments (p.10):** **"Chiao et al. (2009) measured empathic response to stimuli involving pain (and no pain) and found higher SDO scores covaried with reduced activation in the anterior cingulate cortex (ACC) and insula when perceiving pain in others."**
10	**29-31**	I don't follow the reasoning here. It could well be the case that SDO covaries with STS and dlPFC activity when viewing faces of different perceived ranks in real life, regardless of how social rank is defined ? **We thank the reviewer for this clarification and agree that SDO is likely associated with perceived ranks in real life. However, the finding that SDO scores covary with STS and dlPFC is based on a loose operational definition of 'superiority in a social hierarchy'. In the case of Ligneul et al.'s study, this was defined as 'competitive skill' in the task. That is, the authors implicitly defined 'winner' as 'more superior in a social hierarchy'. Because there was no manipulation check of whether participants perceived 'winners' as 'higher in rank' than another in a social hierarchy, whether SDO scores covary with STS and dlPFC activity due to perceived ranks is debatable.** **We have also clarified this in the main text (p.10):** **"It is possible that participants did not perceive any social ranking during the task at all. As there was no manipulation check for this implicit assumption, it is not clear why SDO scores covaried with dlPFC and STS activity. Consequently, the association between SDO and dlPFC and STS regions may not be borne out once this particular task is no longer carried out during the brain scan."**

11	**9**	I don't understand why H4 includes only ACC, and not midcingulate cortex as well? The description of Cazzato et al's findings mentions other regions within the "social orienting circuit" – I'm not sure why the hypothesis is specific to the MCC/ACC & insula, rather than including the whole social orienting circuit? **We thank the reviewer for bringing this clarification to our awareness. Research has made clear distinctions between ACC and MCC and it was a mistake on our part to assume that one could substitute for the other. Given that only the insula was a region that was found to be implicated with SDO in both the Cazzato et al. and Chiao et al. studies, we opted to focus H4 only on the association between SDO and insula. Additionally, the Cazzato et al paper did not make any specific hypothesis of SDO scores covarying with the whole social orienting circuit so we hesitate to make any assumptions. We have also amended this in the main text (p.11):** **"Therefore, we believe it's likely that variation in SDO scores will be negatively associated with structural volume of the insula (H4)."** **We understand that this might cause some misunderstanding so we have now deleted the paragraph in the introduction where we referred to the social orienting circuit.**
11	**15**	The Asp et al study looked at vmPFC lesions, not brain activity? **We thank the reviewer for this clarification and as such, this has been fixed (p.12):** **"Both RWA and SDO has been shown to exhibit cross-cultural validity (Roccato & Ricolfi, 2005), temporal stability (Osborne et al., 2017; Ruffman et al., 2019), and association with brain lesion (RWA: Asp, Ramachandran & Tranel, 2012;) and brain activity (SDO: Cazzato et al., 2015; Chiao et al., 2009)."**
11	**26**	As far as I can tell, there is insufficient grounds to predict that SDO scores will not correlate with vmPFC. As the of Cazzato et al. study used an ROI analysis, and the Chiao et al study reported activation in PFC. **We thank the reviewer for this point and have made the necessary elaboration in the main text to explicate the double dissociation of SDO and RWA (p.11):** **"Our prediction of non-overlapping neuroanatomical regions associated with RWA and SDO suggests an independence of function between these two ideologies at the neural level. Though there is no direct evidence for this double dissociation, some indirect evidence in the literature hints to this possibility. The study by Asp, Ramachandran & Tranel (2012) demonstrated that only damage to vmFPC was**

		significantly associated with higher RWA scores compared to healthy controls. Patients with damage to other neural structures, including those that are involved with emotion, did not show this increase in RWA scores. The etiologies of these non-vmPFC lesions were not overly specified. Notwithstanding, this distinctiveness of RWA scores associated with only vmPFC damage and not other cortical regions implicated in emotional processing leads us to hypothesise that RWA is likely not to associate with the insula. Importantly, this non-vmPFC lesion group excludes patients with specific damage to the amygdala. Thus, the predicted overlapping association of RWA and SDO with the amygdala remains intact. Moreover, only Chiao et al. (2009) has thus far conducted a whole-brain analysis to identify regions that covary with SDO scores during an fMRI task. SDO scores were a significant predictor of frontal areas, namely, inferior, superior and middle frontal gyri activity, in addition to the aforementioned ACC and insula activity when participants engaged in an empathic task. However, after controlling for age and self-reported dispositional empathy, only the ACC and insula were left as regions significantly associated with SDO scores. To our knowledge, no other studies have conducted a whole-brain analysis involving SDO. Comparing with the study by Cazzato et al. (2016), only the insula region consistently covaries with SDO scores across different fMRI tasks. Based on the limited research on this topic, we hypothesise that SDO but not RWA will be associated with the insula and RWA but not SDO will be associated with vmPFC."
11	**27**	Similar to the above point, I'm not sure there's sufficient grounds to hypothesise that RWA scores would not correlate with ACC and insula. **We thank the reviewer for this point and would kindly direct them to the text above.**
14	**27**	Will data be checked to make sure it meets assumptions for parametric analysis before Pearson's correlation is run? **We thank the reviewer for this point and address this in the methods on p.15:** **"Prior to running correlational analysis, RWA and SDO scores will undergo preliminary analysis to check that assumptions for parametric testing are fulfilled. In particular, the normality of responses for each item will be verified using skewness and kurtosis recommended cut-off of +/- 3 values. Items that do not fulfil the normality assumption will be removed from further analysis."**
17	**8**	How will the ROIs be defined? **We thank the reviewer for this clarification and address this point in the methods on p.18:**

		"We will investigate the association between GMV and scores in the RWA and SDO scales using ROI multiple regression analyses. As the brain areas in our hypotheses are identified with a strong a priori prediction, the threshold of significance was set at $p < 0.05$, with small volume correction for multiple comparisons in the ROIs. To analyse these regression models, we will be using the MarsBaR toolbox (http://marsbar.sourceforge.net/) to anatomically define ROIs for the three pre-specified brain regions (i.e. amygdala, vmPFC and insula) according to our main hypotheses. The GMVs will be extracted from their ROIs using anatomically defined spheres with a radius of 20mm centered at (MNI: x=−36, y=−9, z=−17) for the left amygdala and at (MNI: x=27, y=12, z=−21) for the right amygdala (Nam et al., 2017). We then average the mean volumes from the left and right amygdala. The ROI of the vmPFC will be anatomically defined as a sphere with a radius of 20 mm centered at (MNI: x = 0, y = 40, z = -18; Li et al., 2017). Finally, the ROI of in the insula will be anatomically defined as a sphere with a radius of 20 mm centered at (MNI: x = -45, y = 26, z = -6; Chiao et al., 2009; Cazzato et al., 2015)."
17	**14**	Sex should be controlled for as well as age **We agree with the reviewer and we will include age, gender as covariates in the model. As per another reviewer's recommendation, we have also included global brain volume in the form of total intracranial volume as an additional covariate. Please see the detailed description on p.19:** **"In these analyses, we intend to control for total intracranial volume (TIV), age and gender by including them into the regression model as independent "nuisance" variables. TIV is an important variable to account for particularly in ROI-based volumetric measures because such subtle differences in regional brain volume may be confounded by individual differences in overall brain size (O'Brien et al., 2011). We are also controlling for age not only because TIV varies as a function of age (Bartholomeusz et al., 2002), but also because both RWA and SDO have been shown to decrease with age (Altemeyer, 1998; Ruffman et al., 2020). Accounting for age is also necessitated in this study because the analysis will include participants from two different age groups, a young adult sample and a middle-aged adult sample. We would expect both self-report and volumetric brain differences between these two age groups so including age in the regression model will minimise confounds due to age differences. Finally, past research also suggests a gender difference in self-reports of RWA and SDO. In particular, women tend to report higher RWA scores than men (Brandt & Henry, 2012) whereas men tend to report higher SDO scores than women (Pratto et al., 1994). Combined with an overall brain volume differences between men and women (Kaufmann et al., 2001; Ruigrok et al., 2014; Takahashi et al., 2011), we reckoned controlling for gender would facilitate in identifying significant neuroanatomical correlates, as we**

		predict with the age variable. We would like to emphasise that although system-justifying ideologies and regional (and overall) brain volume do seem to vary with age and gender, these are treated as nuisance variables in the main analysis as they do not comprise the main objectives of the study."
18		In the 6th column, it says that a non-significant or significantly negative correlation would disconfirm the hypothesis. However, absence of evidence is not evidence of absence. **We thank the reviewer for this point and have made the change in the manuscript (p.21):** **"Significantly negative correlations [between RWA and SDO] would disconfirm the hypothesis."**

We thank the reviewer for this insightful suggestion and have implemented a combined ROI and whole brain analysis approach to our study. We have added an additional subsection, "2.9. Whole Brain Analysis Plan" (p.20):

"To supplement the a priori ROI analysis, we will also be conducting a whole brain analysis using the DARTEL package in SPM12. As with the ROI analysis, RWA or SDO scores will be used as contrasts to test significance of regressions coefficients from zero value. Similarly, age, gender and TIV will be included as covariates. Significance thresholds will be set at a peak-level threshold of $p < 0.05$ with family-wise error (FWE) correction, and uncorrected voxel-wise level of $p < 0.001$."

Reviewer 3

1A. The scientific validity of the research question(s).

The authors investigate neuroanatomical correlates of system-justifying ideologies, i.e., trait measures of social dominance (SDO) and right-wing authoritarianism (RWA). This research question is scientifically justifiable and derived from existing theories. Furthermore, it is defined with sufficient precision as to be answerable through quantitative research. It also falls within established ethical norms.

1B. The logic, rationale, and plausibility of the proposed hypotheses, as applicable.

The authors state clear, directional hypotheses. Yet, there is some additional literature that seems to be relevant for deriving the hypotheses (e.g. Baumgartner et al., 2014, Neuroimage, “Impartiality in humans is predicted by brain structure of dorsomedial prefrontal cortex”). While I found the direction of these hypotheses reasonably deduced from previous literature, I was a bit confused regarding the underlying unidimensionality and the expected correlation between SDO and RWA. If the correlation among these measures depends so heavily on the cultural/political environment, is it then plausible to expect a common evolutionary/biological basis of these two constructs? Maybe the authors could clarify this issue.

We would like to thank the reviewer for raising this insightful reference to our awareness. However, we restricted our literature search to derive our hypotheses on neuroanatomical correlates using references that included either RWA or SDO (or in the case of Nam et al., system-justification scale, which is of relevance to our framework using system-justifying ideologies) in their neuroimaging study. Moreover, our study did not include any measure of impartiality and so hesitate to include this specific reference.

We would also like to thank the reviewer for highlighting this and as such, we have changed the section header to “1.5. Right Wing Authoritarianism and Social Dominance Orientation as Separate Constructs” (p.6) to minimise confusion regarding the unidimensionality and expected correlation between RWA and SDO. We would also like to further clarify how the correlation between SDO and RWA may relate to the biological basis of the constructs:

Although the correlation between RWA and SDO scores vary as a function of the degree of ideological contrast of the population of study, this does not preclude the distinction of antecedents of RWA and SDO as predicted by the DPM model. That is, RWA is largely motivated by a dangerous worldview whereas SDO is largely motivated by a competitive worldview. In addition, the amygdala is also involved as this region is implicated with legitimising social hierarchies, both promoted by RWA and SDO ideologies. In other words, the degree of ideological contrast has more to do with the correlation between the two ideologies and much less the underlying psychological and neuroanatomical mechanisms underlying each of them. Theoretically, regardless of the degree of ideological contrast of a given population, the shared and distinct brain regions of RWA and SDO would remain stable. That being said, a cross-national comparison of RWA and SDO scores and their neuroanatomical correlates between countries of different ideological contrasts must be conducted to directly address this question.

1C. The soundness and feasibility of the methodology and analysis pipeline (including statistical power analysis or alternative sampling plans where applicable).

In total, I found the methodology and analysis pipeline sound and feasible. The authors have provided a reasonable justification for choosing this sample size. However, I think the authors should add some details to their analysis plan (see 1D).

1D. Whether the clarity and degree of methodological detail is sufficient to closely replicate the proposed study procedures and analysis pipeline and to prevent undisclosed flexibility in the procedures and analyses.

I was missing some details regarding the ROI-definition and statistical analysis. How exactly will the ROIs be defined (e.g. which atlas will be used)? How do the authors plan to test their hypotheses? Could they please specify how they will perform their "ROI multiple regression analysis"? Do the authors expect differences between gender? Why not include gender as an additional covariate? How exactly will the GMV be calculated (e.g., which software will be used)?

We address the point on defining ROI and calculating GMV in the methods on p.18:

''We will investigate the association between GMV and scores in the RWA and SDO scales using ROI multiple regression analyses. As the brain areas in our hypotheses are identified with a strong a priori prediction, the threshold of significance was set at $p < 0.05$, with small volume correction for multiple comparisons in the ROIs. To analyse these regression models, we will be using the MarsBaR toolbox (http://marsbar.sourceforge.net/) to anatomically define ROIs for the three pre-specified brain regions (i.e. amygdala, vmPFC and insula) according to our main hypotheses. The GMVs will be extracted from their ROIs using anatomically defined spheres with a radius of 20mm centered at (MNI: x=−36, y=−9, z=−17) for the left amygdala and at (MNI: x=27, y=12, z=−21) for the right amygdala (Nam et al., 2017). We then average the mean volumes from the left and right amygdala masks. The ROI of the vmPFC will be anatomically defined as a sphere with a radius of 20 mm centered at (MNI: x = 0, y = 40, z = -18; Li et al., 2017). Finally, the ROI of the insula will be anatomically defined as a sphere with a radius of 20 mm centered at (MNI: x = -45, y = 26, z = -6; Chiao et al., 2009; Cazzato et al., 2015)."

We thank the reviewer for their recommendation and have expanded on how the multiple regression analyses will be conducted on p.19:

"Each ROI multiple regression analysis will use ordinary least squares models with the GMV as the dependent variable, and RWA (or SDO) score, gender, age and TIV as independent variables."

We agree with the reviewer and we will include gender as covariates in the model. As per another reviewer's recommendation, we have also included global brain volume in the form of total intracranial volume as an additional covariate. Please see the detailed description on p.19:

''In these analyses, we intend to control for total intracranial volume (TIV), age and gender by including them into the regression model as independent "nuisance" variables. TIV is an important variable to account for particularly in ROI-based volumetric measures because such subtle differences in regional brain volume may be confounded by individual differences in overall brain size (O'Brien et al., 2011). We are also controlling for age not only because TIV varies as a function of age (Bartholomeusz et al., 2002), but also because both RWA and SDO have been shown to decrease with age (Altemeyer, 1998; Ruffman et al., 2020). Accounting for age is also necessitated in this study because the analysis will include participants from two different age groups, a young adult sample and a middle-aged adult sample. We would expect both self-report and volumetric brain differences between these two age groups so including age in the regression model will minimise confounds due to age differences. Finally, past research also suggests a gender difference in self-reports of RWA and SDO. In particular, women tend to report higher RWA scores than men (Brandt & Henry, 2012) whereas men tend to report higher SDO scores than women (Pratto et al., 1994). Combined with an overall brain volume difference between men and women (Kaufmann et al., 2001; Ruigrok et al., 2014; Takahashi et al., 2011), we reckoned controlling for gender would facilitate in identifying significant neuroanatomical correlates, as we predict with the age variable. We would like to emphasise that although system-justifying ideologies and regional (and overall)

brain volume do seem to vary with age and gender, these are treated as nuisance variables in the main analysis as they do not comprise the main objectives of the study.”

1E. Whether the authors have considered sufficient outcome-neutral conditions (e.g. absence of floor or ceiling effects; positive controls; other quality checks) for ensuring that the obtained results are able to test the stated hypotheses or answer the stated research question(s).

Not applicable.

In closing, we would like to thank all reviewers for their careful reading and all the suggestions for how we can improve our work. We hope all reviewers now find our registered report Stage 1 conclusive, interesting and ready for Stage 2.

Best Regards,

Prof Esposito and colleagues

Associate Professor

Social & Affective Neuroscience Lab Psychology Program - SSS, Nanyang Technological University, Singapore

Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore

Affiliative Behaviour and Physiology Lab Department of Psychology and Cognitive Science, University of Trento, Italy