Social cognition as a matter of structural brain connections: a systematic review and diffusion weighted imaging meta-analysis

Rita Hansl, Lara Z. Maliske, Sofie Valk, Philipp Kanske

OSF project folder: https://osf.io/3z4bf/?view_only=ca95cb2546604b6ab7da562fbee68d39
Stage 1 manuscript: https://osf.io/bemw7?view_only=ca95cb2546604b6ab7da562fbee68d39
ArticleID #907
version 2

Recommender: Marietta Papadatou-Pastou

Decision for round #2 : Revision needed - Minor revision

The reviewers and I have now evaluated the revised manuscript. One reviewer and I found the revisions satisfactory. However, the second reviewer still requires a few changes. I think it is important to implement them before in-principle acceptance of Stage 1 because the authors will need to implement the suggestions in their meta-analysis. I am looking forward to the revised manuscript, which I expect will not take too long for the authors to prepare.	Thank you very much again for the smooth handling process and the very constructive and respectful feedback. We appreciate the input a lot and are happy to improve the manuscript further. We hope we could accomplish this improvement to a satisfying degree for all reviewers in this second round and look forward to your decision.

Reviewer 1: Sebastian Ocklenburg

I have read the response to reviews letter and the marked revised manuscript and can confirm that the authors improved an already strong RR even further. I have no further suggestion and can recommend that the RR can be accepted by PCI.	Thank you very much for this very supportive and kind feedback. We are happy to hear that you find our work satisfactory and want to thank you again for the constructive input in review round 1.

Reviewer 2: Katie Lavigne

I would like to thank the authors for carefully integrating my comments and suggestions. I only have a few concerns remaining:	Thank you very much for the further examination of our manuscript as well as the constructive input.
Point m (as per author reply): Thank you for clarifying the coding procedure. However, I would like to re-iterate my point about the "score" as being separate from the measure. It is possible that a measure provides multiple scores (e.g., subscales for questionnaires or both accuracy and reaction time for tasks) that may capture different socio-cognitive constructs or capture the same ones in a different way. This might require different decisions with	Thank you very much for clarifying this point and eliminating our misunderstanding. We agree that it is essential to record the specific score/subscale and not exclusively the measurement tool. We made this point more explicit in the manuscript and integrated another column in our data recording sheet.

regard to construct classification or analysis (e.g., reverse coding) that would be important to note during data extraction.	
Point r: I would recommend the authors consider excluding findings that are not standard correlations, or conduct sensitivity analyses, as these effects will likely complicate the meta-analytic findings and increase heterogeneity. For example, with multivariate techniques, the effects will not be comparable to standard correlations as the values integrate the effects of many other variables. This is quite different from a standard correlation and is highly dependent on the model. Partial correlations raise a similar concern, but could be integrated if clear patterns are identified (e.g., age & sex being partialled out) as with potential sensitivity analyses.	Thank you very much for raising this important point. We discussed the topic in detail and suggest the following approach: Given the relevant literature reviewed by (Wang et al., 2018) we expect authors of included studies to perform correlation as well as regression analysis. Based on suggestions by authors such as (Nieminen, 2022), standard regression coefficients will be calculated from the different reported effect sizes. Where available, main effects of regression models will be reported as effect sizes. However, we expect relevant studies to perform multivariate regression analysis and would like to be as inclusive as possible, while assuring comparability. Therefore, we added another column to the coding sheet to report the type of analysis performed in the study as well as the variables controlled for in the model corresponding to the reported effect size. We agree that sensitivity analysis is important and have therefore already planned various approaches in the manuscript. These are discussed in section ***2.7.5 Publication bias*** which has been renamed to ***2.7.5 Publication bias and sensitivity analysis***. The planned types of sensitivity analysis include: - Funnel plot - Egger's test - PET & PEESE model in Robust Bayesian meta-analysis - Jackknife/Leave one out sensitivity analysis in the coordinate based meta-analysis (section ***2.7.3 MA2***)
Point t: Please add "any" to specify "failing to report relevant details on any of the defined moderators...".	Thank you for this remark. We implemented the suggested changes.
Addition: The addition of a measure of inter-rater reliability is welcomed, but I would like to caution the authors on the use of Cohen's Kappa, which can produce low values despite high agreement (the Kappa Paradox, https://doi.org/10.1016/j.jhsa.2024.01.006) with systematic reviews, due to the discrepant prevalences of include/exclude ratings. I would recommend Gwet's AC1 statistic or the Brennan-Prediger coefficient, as these performed well in our previous scoping review (https://doi.org/10.1038/s41537-022-00219-x, Supplementary Table 1).	Thank you very much for the suggestion and cautioning us for this issue. We read into the pros and cons of either measure. Based on the thorough discussion with our statistician and a review of relevant literature, we decided for the use of Cohen's Kappa for the following reasons: - Despite the highlighted disadvantages, the measure is the most established one and hence also most likely to be interpreted correctly by readers. - In our opinion, the key issue does not lie with the measure itself but rather in its interpretation. Thanks to your input and the

	subsequent elaboration on the topic we are now more aware of the issue and will ensure appropriate discussion and interpretation of the results (Derksen et al., 2024). Importantly, the expected agreement by chance will be reported to facilitate appropriate interpretation, even if a seemingly low value were to result. - Alternative measures may yield higher values for agreement. However, they are also more complex and less intuitive for us (and hence maybe also potential readers) to interpret. However, we also appreciate the benefits of the newer methods (Wongpakaran et al., 2013). We therefore commit to additionally reporting Gwet's AC1 to provide a measure for both agreement and disagreement between raters. This has been added in the manuscript in section ***2.5 Screening procedure.***